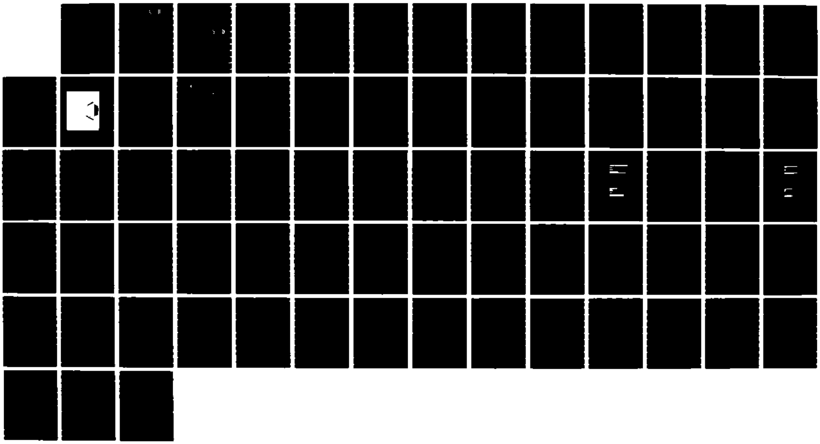
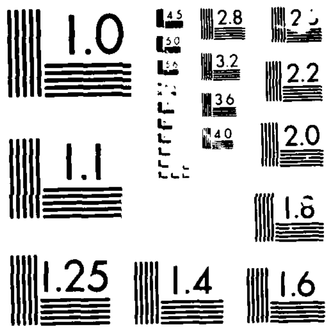
AD-A169 514 SIMULATOR DESIGN FEATURES FOR HELICOPTER LANDING ON SMALL SHIPS I A PERFO..(U) ESSEX CORP ORLANDO FL D P WESTRA ET AL. 27 SEP 85 NAVTRASYSCEN-81-C-0105-13

UNCLASSIFIED N61339-81-C-0105 F/G 5/9 NL

1/1




1.0
2.8
2.2
3.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

(12)

HELICOPTER LANDING ON SMALL SHIPS.
I. A PERFORMANCE STUDY

Daniel P. Westra and Gavan Lintern

Submitted by:

Essex Corporation
1040 Woodcock Road
Orlando, Florida 32803

Prepared for:

NAVTRASYSCEN 81-C-0105-13

12

SIMULATOR DESIGN FEATURES FOR
HELICOPTER LANDING ON SMALL SHIPS.
I. A PERFORMANCE STUDY

Daniel P. Westra and Gavan Lintern

Submitted by:

Essex Corporation
1040 Woodcock Road
Orlando, Florida 32803

Prepared for:

Naval Training Systems Center
Orlando, Florida 32813

Contract No. N61339-81-C-0105

Interim Final Report for the Period
1 September 1981 - 26 April 1985

DOD DISTRIBUTION STATEMENT

Approved for public release,
distribution unlimited

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A169514

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) NAVTRASYSCEN 81-C-0105-13
6a. NAME OF PERFORMING ORGANIZATION Essex Corporation — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Naval Training Systems Center
6c. ADDRESS (City, State, and ZIP Code) 1040 Woodcock Road Orlando, FL 32803	7b. ADDRESS (City, State, and ZIP Code) Orlando, FL 32813
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Naval Air Systems Command — 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N61339-81-C-0105
8c. ADDRESS (City, State, and ZIP Code) Washington, DC 20361	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
N637-33N	SP-01	0785-2P3	4780

11. TITLE (Include Security Classification)
Simulator Design Features for Helicopter Landing on Small Ships. I. A Performance Study

12. PERSONAL AUTHOR(S)
D. P. Westra and G. Lintern

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim Final	FROM 9/81 TO 4/85	27 September 1985	69

16. SUPPLEMENTARY NOTATION

Keywords:

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
05	08	
14	02	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Flight simulator, visual simulation, scene content, helicopter landing, landing at sea, simulator design, ~~field of view~~

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
The Visual Technology Research Simulator (VTRS) at the Naval Training Systems Center was used to study the effects of six simulator features on performance for helicopter landings on small ships. The purpose of the experiment was to obtain information relevant to the design of simulators for skill maintenance and transition training, and to obtain information for decisions about future transfer-of-training studies. The six simulator factors were scene detail (high detail ship deck and hangar markings versus no deck and hangar markings), field of view (VTRS-wide versus reduced SH-60B operational flight trainer field of view), system visual lag (217 msec versus 117 msec), g-seat acceleration cuing (off versus on), g-seat vibration cuing (off versus on), and collective sound cuing (off versus on). These factors were tested across two levels of

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Dennis C. Wightman	(305) 646-5130	711

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

19. Abstract (cont'd)

seastate and pilot experience. Pilots who participated in the experiment were experienced Navy H-3 rotary wing pilots. Results indicated large effects of scene detail, small to moderate effects for visual lag, small effects for field of view, and no meaningful effects for the g-seat factors and collective sound. Performance was better with the high-detail ship, the shorter visual lag, and the VTRS-wide field of view. Transfer-of-training research is recommended as the next step to further explore these findings and to obtain information directly relevant to the design of simulators for use by student pilots.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ACKNOWLEDGEMENTS

The entire VTRS on-site technical staff deserves special mention for providing the support necessary to conduct this experiment. In addition, we wish to thank LCDR Larry Cable, USN; LCDR Dave Crocker, USN; LT Wayne Nelms, USN; and LT James Thompson, USN, who served as advisors in setting up the experimental task. We also wish to thank the pilots from NAS Jacksonville and NAS Mayport who volunteered to take part in the experiment. Finally, special credit goes to Ronald Mauk who performed much of the performance measurement work and data analysis.

SECTION I

INTRODUCTION

Recent years have seen a marked expansion of the role of helicopters in the US Navy and, in particular, the requirements for operations from small ships has expanded significantly. Helicopters now participate in the roles of surveillance, communications, resupply, and attack. Many Navy vessels, both large and small, support helicopters and most have a helipad constructed for that purpose. While aircraft operations at sea generally pose additional hazards over those of land-based operations, the problems are accentuated on small ships because of the confined landing area and because of instability of the deck. This report describes an experiment that initiates research at the Naval Training Equipment Center's Visual Technology Research Simulator (VTRS) into simulator design and use in teaching essential flight control skills for helicopter operations from small ships.

The simulation at the VTRS facility supporting helicopter training research includes an SH-60B cockpit with all displays and controls that are important for flight control and guidance. These displays and controls function in real time and closely simulate those of the aircraft within the flight regime of the approach and landing. The cockpit is mounted on a fixed base in a 17-foot (5.18m) radius dome. It has a pneumatic g-seat, with buttock, thigh, and back cushions that simulate tactile pressures experienced in flight. Twin 1025-line color projectors are used to provide a 160 degree (H) by 70 degree (V) computer-generated image of the outside visual scene. The field of view is 40 degrees to the left, 120 degrees to the right, 20 degrees above, and 50 degrees below the forward line of sight. Maximum scene brightness is approximately 0.2 ft Lamberts (0.685 cd/m^2). Herndon (1982) provides a more complete description of the VTRS helicopter simulator.

PROBLEM

Landing a helicopter on a small ship is a particularly difficult task and that difficulty is accentuated in turbulent seas. Typically, the pilot establishes the aircraft on a descent path about one mile behind the ship and approaches the landing area while reducing speed. The pilot arrests the descent near the stern of the ship and taxis over the landing area at a height of approximately 15 feet. A hover is maintained above the touchdown point until the pilot ascertains that the deck is level and stable enough for a safe landing. At that moment, the aircraft is quickly lowered to the landing area and secured to the deck. Simulators are available for

some aspects of ship-based operations, but an informal survey during planning for this experiment indicated that none is satisfactory for teaching the final stages of the helicopter approach and landing.

The Light Airborne Multipurpose System Mark III (LAMPS MK III) integrates an FFG7 frigate and a SH-60B Sea Hawk helicopter to provide an over-the-horizon detection and strike capability for antisubmarine warfare and antiship surveillance and targeting. The system is currently being introduced to the US Navy and it is anticipated that approximately 100 units (i.e., 100 ships and 200 aircraft) will eventually be deployed. A simulator that would support the LAMPS MK III system, by enabling instruction in flight control for ship-related operations, is desired. The SH-60B cockpit was installed at the VTRS facility so that simulator design and instructional issues for small ship operations could be examined.

The experiment reported here provides initial design guidelines adequate for teaching the most critical skill (that being shipboard landings), and provides information for subsequent research that will examine design and instructional issues in more detail.

RESEARCH APPROACH

The experiment was designed as a performance study in which pilots who could perform the task with at least reasonable proficiency in both the aircraft and the simulator were tested on a variety of experimental conditions to provide information about absolute and relative effects of the experimental factors. This type of study should be distinguished from a transfer study in which subjects who are not proficient in the designated task are taught it under one of several possible experimental conditions, and are then tested on a criterion condition.

Past and ongoing research at the VTRS into simulator design for flight training has employed a three-phase approach. Performance studies of the type described above have constituted only the first phase. The second phase involves quasi-transfer experiments in which both training and transfer are conducted in the simulator. The highest fidelity configuration of the simulator is used for the transfer or criterion condition. The third phase involves a simulator-to-field transfer experiment in which the effects of prior training in a simulator are tested in an airplane. Information obtained in the early phase is used in planning the experimental strategy for succeeding phases. As the conduct of each successive phase is considerably more costly, this strategy helps ensure that information for simulator design will be obtained in a cost-effective manner.

The three-phase approach has been used in VTRS research to determine simulator design requirements for a carrier landing trainer (Westra, Simon, Collyer, and Chambers, 1982; Westra, 1982, and an air-to-ground attack trainer (Westra, 1983; Lintern, Thomley, Nelson, and Roscoe, 1984. This research has been summarized by Lintern, Wightman, and Westra (1984).

The performance experiment reported here had three goals. The most urgent was to specify design guidelines for a simulator that could be used to teach and maintain flight skills required for SH-60B operations in proximity to a host ship. It could be argued that data from a performance study should not be used to make inferences about training requirements. Reviews of research by Lintern and Gopher (1978), and Wightman and Lintern (1984), have shown that performance effects resulting from the use of different conditions in a training phase often do not result in different levels of learning as shown by a test on a criterion condition. Nevertheless, good training experiments are difficult to conduct and are time-consuming, and research at the VTRS has shown that they should be preceded by considerable developmental work, including at least one performance study. In view of the urgency of the problem and the lack of better data, this performance study was oriented towards providing a preliminary guide to simulator design. We recognize that this may not produce the optimum design, either in terms of learning or of cost effectiveness, but it can be expected to result in the design of an effective learning environment.

A second goal of the performance experiment was to screen variables for subsequent transfer and quasi-transfer studies. Factors that have little effect in this first performance study may be excluded from later transfer studies. The implicit assumption in this screening procedure is that there will be no worthwhile differential transfer effect from a manipulation that does not result in a substantial performance effect. This assumption can be questioned, but exceptions are difficult to find in the experimental literature. In reality, preselection of variables has always been a part of the planning for transfer studies and, in our case, where theory and prior data do not generally offer a useful guide, the results of a performance study are more useful than the alternative of subjective analysis.

A third important goal was to develop and validate performance measures and experimental procedures for subsequent transfer research. One element of this third goal was to identify, via thorough examination of the data, specific difficulties that pilots might have with the task. The identification of such difficulties might suggest instructional strategies or visual guidance systems that could be tested in later experiments.

Six simulator design features were examined. They were field of view, visual system lag, scene detail, g-seat acceleration cuing, g-seat vibration cuing, and sound cuing to rotor speed variations resulting from collective inputs. These represented all of the simulator design issues of interest that could be investigated under the current capabilities of the VTRS. In keeping with the screening intent of the experiment (Simon, 1977; 1979), these factors were tested at two levels each, representing high- and low-fidelity options. To extend the generalizability of results, all factors were tested across two markedly different levels of seastate/turbulence and pilot experience.

SECTION II

METHOD

Eight experienced Navy pilots made approaches to and landings on a representation of an FFG7 frigate in the Vertical Take-off and Landing (VTOL) simulation of the Visual Technology Research Simulator (VTRS) facility. The pilots were from operational squadrons and routinely flew VTOL aircraft in helicopter/ship operations. Their current aircraft (the H-3 helicopter or some variant of it) had noticeably different controls, displays, and dynamic response to the Sea Hawk (SH-60B) simulated in the VTRS. However, the VTRS task was similar in most crucial respects to the landing task as flown in the H-3, and these pilots appeared to experience little difficulty in adapting to this new situation. In 1983 the SH-60B had not been deployed and there were insufficient SH-60B experimental pilots available to participate in this experiment.

APPARATUS

The VTOL simulator has a fully instrumented cockpit and all major controls. Visual, aerodynamic, and motion computations are performed at a 30 Hz iteration rate by a SEL 32/77 computer system of high-speed multiple processors. Cyclic, collective, and directional pedal loading is provided by a variable-force control loading system. A g-seat for the pilot, designed to simulate haptic sensors, provides limited maneuver and disturbance motion cues. Four penumatically driven panels in the seat and back pans are synchronized to provide sustained simulation of acceleration vectors. An additional single panel over the seat pan area enhances heave acceleration cues through vertical displacement. Aircraft and environmental sounds are also simulated.

The outside visual scene is represented by computer-generated images that are projected onto a 17-foot radius screen. A General Electric Compu-Scene I (upgraded with extra edge capacity and a distortion correction capability) and a PDP11/55 computer are used to provide a 6000-edge capacity. Two TV light valve color projectors (1025 lines) are used to display the imagery in adjacent fields to give a 160 degree horizontal (40 degrees left to 120 degrees right) by 70 degrees vertical (20 degrees up to 50 degrees down) field of view.

TASK

The experimental task involved the approach and landing of the simulated SH-60B helicopter to a simulated FFG7 frigate moving forward at 10 knots (18.5 Km/h). A depiction of the simulated ship showing the prominent deck markings is shown in Figure 1. The simulated aircraft was initialized at 160 feet (48.8m) altitude on an approach heading 2000 feet (609.3m) behind the ship. The simulated aircraft was initialized at an

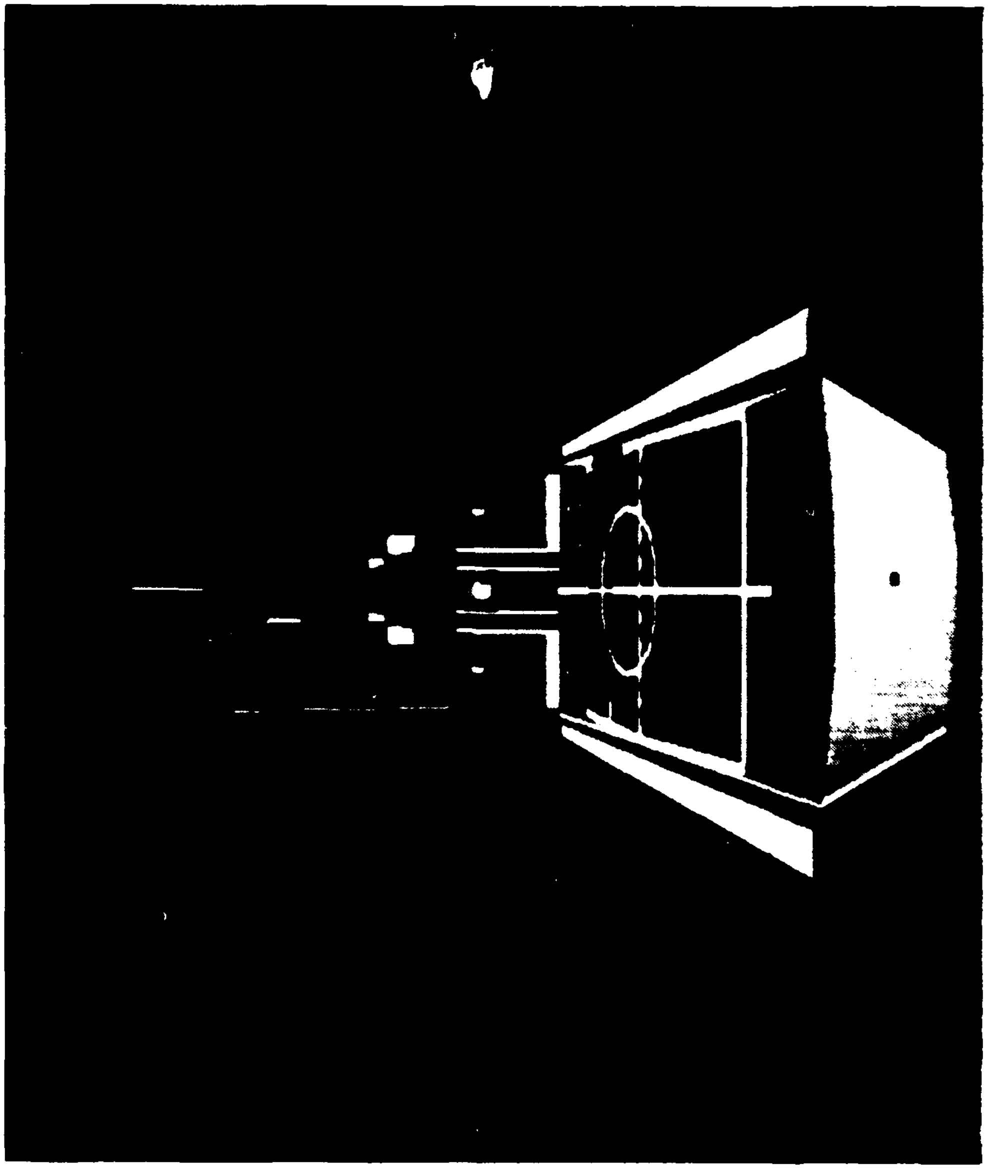


Figure 1. View of Simulated FFG7 Frigate Showing Prominent Deck Markings

airspeed of 43 knots (79.9 Km/h) for a descent rate of 128 fpm (39.0m/min) descending approach to the FFG7. The glideslope approach angle was nominally set at 3.5 degrees, although no glideslope indicator was available and pilots were not specifically concerned with maintaining that approach angle. The approach and landing involved a descending and decelerating approach to the ship, transition to hover near the stern, hover over the landing area, and descent to the designated rapid securing device (RSD).

In fleet operations the approach to landing is flown from directly astern the ship and along the glideslope until the aircraft is within approximately 10 feet (3.05m) of the ship's stern. At this point the descent rate is arrested and the aircraft is maneuvered across the stern to a point approximately 15 feet (4.57m) above the designated landing spot. The pilot adjusts aircraft speed to match that of the ship and a hover is maintained until pitch, roll, and yaw of the ship are momentarily stable and the pilot is prepared to land. The pilot then descends rapidly to the landing deck and the aircraft is secured. A side view depicting an approach and landing is given in Figure 2. For this experiment, our pilot subjects were instructed to follow fleet operational procedures as closely as possible. Further detail is given in Appendix A, which is a copy of the briefing material given to pilots as part of their preliminary instruction.

FACTORS AND LEVELS

Factor level settings were chosen in order to bracket the reasonable range of interest. For equipment factors, high levels were set at the highest fidelity attainable under VTRS capabilities. Low levels were set at the most degraded form of the factor likely to be used in a flight trainer, or at a level currently being used in an operational flight trainer. In some cases, these contrasts also represented considerable cost differences. The seastate/turbulence and pilot experience factors were added to enhance the generalizability of the experiment. Factors and levels are summarized in Table 1.

SCENE DETAIL. The high-detail ship which is depicted in Figure 1 included all prominent deck and hangar markings plus some deck markings intended to represent pad eyes (tie-down fixtures embedded in the deck). A ship's wake was also present, and irregularly shaped, light blue patches were visible on the seascape. The low-detail ship was made up of solid surfaces but had no deck or hangar markings. The shading differences between vertical and horizontal surfaces allowed a perception of the image and permitted discrimination of the landing deck from the bulkhead. No ship's wake was present, and the seascape was featureless. A horizon was available with both the high- and low-detail scenes.

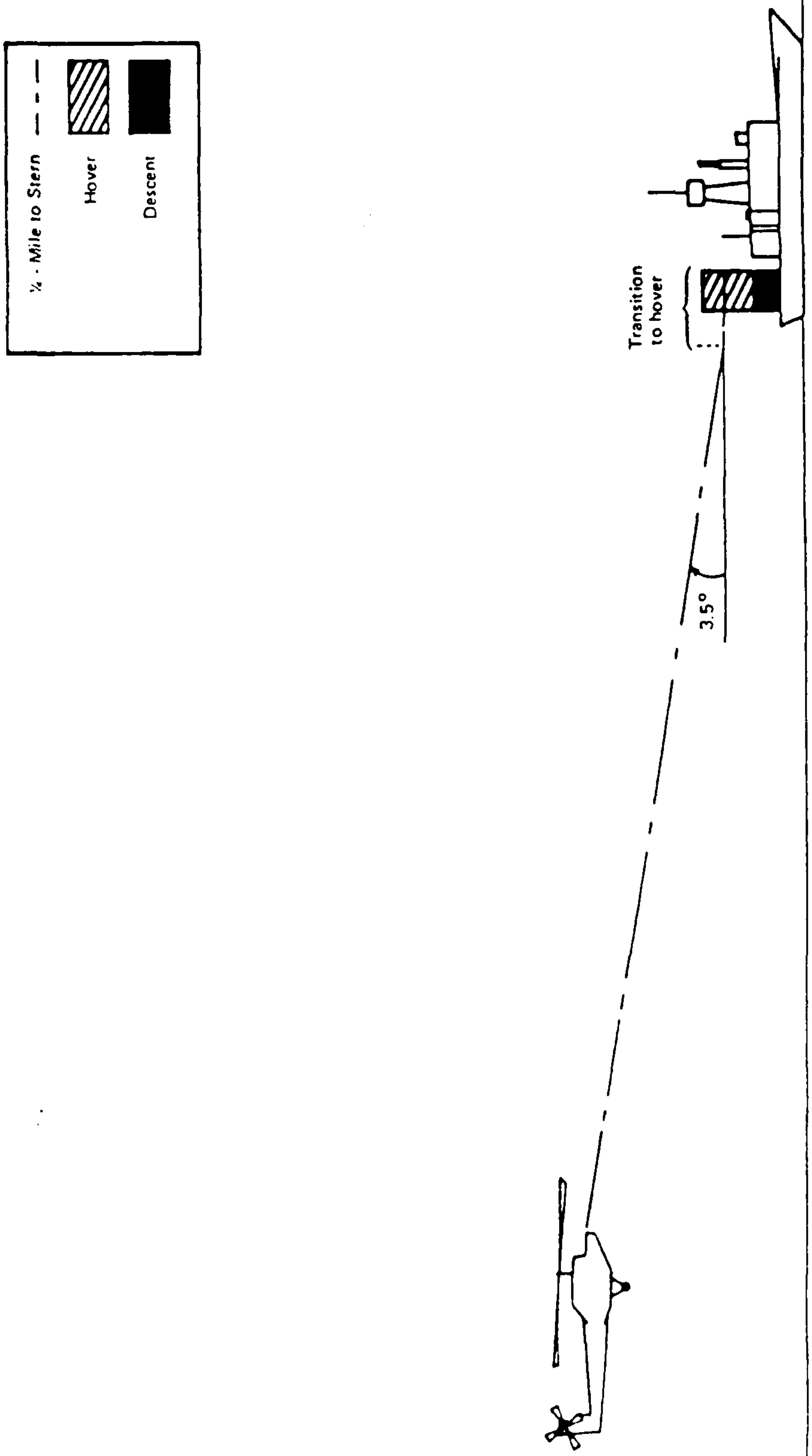


Figure 2. Side view of approach and landing.

TABLE 1. EXPERIMENTAL FACTORS AND LEVELS

FACTOR	LEVELS	
Scene detail	Outline of deck & hangar	Full deck & hangar markings, ship's wake, seascape patterns
Field of view	20^{0} left to 100^{0} right, 15^{0} up to 25^{0} down (left half of field), and 3^{0} up to 40^{0} down (right half of field)	40^{0} left to 120^{0} right, 20^{0} up to 50^{0} down
Visual lag	217 msec	117 msec
G-seat cuing	Off	Translational and angular accelerations
G-seat vibration	Off	Oscillating cushions
Collective sound	Off	Augmented aural cues
Environment	Moderate seastate and medium air turbulence	Calm with no air turbulence
Average flight experience	830 hours	2323 hours

FIELD OF VIEW. The maximum field of view displayed on the dome surface with respect to the pilot's eye was 40 degrees left to 120 degrees right, and 50 degrees down to 20 degrees up (Figure 3). It filled the available windscreen area except for the lower half of the chin window which was beyond the lower limits of the projection system. The visual scene was viewed through the pilot's forward, right, and chin windows. The other cockpit windows, which did not appreciably involve the pilot's field of view, were covered for this experiment.

The reduced field of view resembled the one available in the Navy's SH-60B Operational Flight Trainer. This field of view consisted of two adjoining rectangles with the one in front of the pilot extending 20 degrees left to 55 degrees right and 15 degrees up to 25 degrees down. The right rectangle extended 55 degrees right to 100 degrees right, and 3 degrees up to 40 degrees down with respect to the pilot's eye-point (Figure 3). The downward field of view is a critical area in that the pilot gains considerable information from it during the hover. The reduced field of view provided downward visibility only on the lower right-hand side window and in the upper half of the chin window. Our pre-experimental pilots had

believed that the difference in coverage in the chin window area (see Figure 3) would affect hover performance.

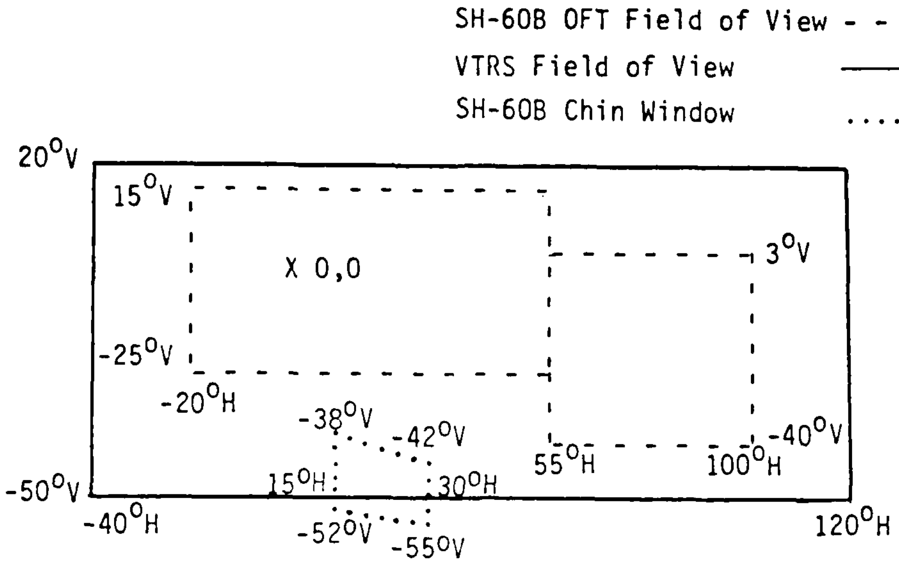


Figure 3. Experimental levels for field of view.

VISUAL SYSTEM LAG. The shorter lag condition averaged 117 msec (100 to 133 msec depending on the moment of stick sample) from stick input to the completion of the first field of video output. This represented a 30 Hz aerodynamic update rate and a 60 Hz CIG update rate (Browder & Butrimas, 1981), and is a better response than is available on most Navy trainers. The longer lag average of 217 msec (199 to 232) represents a 15 Hz aerodynamic update rate and 30 Hz CIG update rate, and produces the slowest response normally considered in acquisitions of simulators with visual systems. The faster response obviously requires greater computer capacity.

G-SEAT ACCELERATION CUING. The pneumatic cushion g-seat (Bose, 1980) was fully operational for the high-fidelity level of this factor. The cushions were driven to represent aircraft translational and angular acceleration cues in the vertical, lateral, longitudinal, yaw, pitch, and roll dimensions. These accelerations were scaled by SH-60B test pilots during pre-experimental work and, because accelerations to be found in actual flight produce only slight haptic sensations, were augmented to provide haptic sensations that exceeded those experienced in flight. The g-seat cuing was not operational for the low-fidelity level of this factor.

G-SEAT VIBRATION. G-seat vibration was manipulated independently of acceleration cuing. Vibration cues representing nominal helicopter background vibration and translational lift were provided by oscillations in the g-seat pan cushion. The frequency of the background vibration was 10 Hz, which was the maximum frequency to give a reasonable amplitude of vibration as defined by SH-60B test pilots. An additional translational lift vibration or buffet was implemented as a low-frequency, large-amplitude vibration which occurred whenever the aircraft decelerated through airspeeds between 25 to 20 knots. No vibration cuing was present for the low-fidelity level of this factor.

COLLECTIVE SOUND. Aural cues were augmented by increasing the amplitude of main rotor blade, main gear box, and auxiliary gear box sounds as a function of collective position and vertical acceleration. This cue was scaled to provide a small increase in sound amplitude with increasing collective to simulate rotor loading effects. No augmented collective sound cuing was present with the alternate level of this factor. Other gear box and main rotor blade sounds were present at all times.

SEASTATE/TURBULENCE. At one level of this factor there was no ship motion (other than the 10-knot forward movement) and no air turbulence acting on the aircraft. At the other level there was ship motion in roll, pitch, yaw, sway, and heave that corresponded approximately to effects expected from moderate seas. Typical RMS values of movement were 3.2 degrees, 0.5 degrees, 0.4 degrees, 0.4 feet (0.12m), and 0.4 feet (0.12m) for roll, pitch, yaw, sway, and heave, respectively. This amount of ship activity might be produced for waves averaging 1.8 feet (0.55m) in height. Complete details of the ship dynamics and air turbulence models are contained in VTRS software modules. With an active seastate there was also air turbulence acting on the aircraft's vertical, lateral, and longitudinal axes corresponding approximately to that expected with the given seastate.

PILOT EXPERIENCE. All subjects in the experiment were experienced helicopter pilots. Thus, the levels for this factor represent a relative difference within an experienced group. Pilots categorized as less experienced had an average of 830 rotary wing flight hours (range of 400 to 1250 hours). Pilots categorized as more experienced averaged 2323 flight hours (1500 to 3000 hours).

PERFORMANCE MEASUREMENT

Raw data were recorded at 30 Hz and were reduced to a set of trial-summary measures. For measurement purposes, the task was partitioned into four segments, those being: 1) an approach from 1500 feet astern to the stern of the ship, 2) transition from the stern to a hover above the landing point, 3) hover

above the landing point, and 4) descent to touchdown (see Figure 2). Each pilot was set up on glideslope and on centerline 2000 feet astern. He was then given control of the aircraft and was required to descend to the ship and to land.

Position and velocities of the simulated aircraft were recorded throughout the exercise. The approach, transition, hover, and descent segments over which data were summarized are described in Table 2. Some problems were experienced with hover scores in that pilots were occasionally so low when they crossed the stern that they were never in the hover segment. Hover scores were considered missing if less than 2 seconds of hover time was recorded. A total of 80 summary measures were computed on-line for each trial. The entire list of summary measures collected on-line is given in Appendix B. Some measures such as glideslope RMS error in the descent have little or no meaning and were not used in analyses.

TABLE 2. TASK SEGMENT DEFINITIONS FOR EXPERIMENTAL TASK

	Definitions of Task Segments
Approach	The segment started 1500 feet astern and continued until the center of gravity of the simulated aircraft crossed the stern of the ship.
Hover	This segment commenced 2 seconds after the aircraft center of gravity came within 3 feet (0.91m) of a point directly above its optimum position at touch down. In those trials in which the aircraft crossed the stern at more than 13 feet (3.96m) above the deck, the segment continued until the aircraft descended to less than 9 feet (2.74m) above the landing deck. When the aircraft crossed the stern of the ship at less than 13 feet above the landing deck, it continued until the aircraft descended below 70% of the stern-crossing height.
Descent	This segment commenced at the end of hover and continued until weight was recorded on all wheels for at least 3 seconds.

A large number of summary measures are available, and it was possible to create others from the raw data. These were categorized and prioritized to ensure that the data would be presented in a manner that would aid interpretation of the results. The measures may be thought of as belonging to the pilot-aircraft control loop and can be broadly categorized as pilot input variables, aircraft output variables, and task outcome variables. Because pilot input variables, such as

collective movement, are typically highly pilot-dependent, they were not given much weight in data analysis interpretation. Aircraft output variables, such as roll variability, may reflect task difficulty but may not be directly associated with visible indices of accurate or correct performance. Task outcome variables directly measure aircraft position relative to the optimum or desired flight path as defined by the task. In general, effects that were reflected in outcome variables are considered the most important.

A performance measure may be: (1) an integration of a variable over a segment of time, or (2) a single value or snapshot taken at a specific point. Measures that are integrated over a segment provide a within-trial summary. Any given variable, such as lineup deviation during approach, may be integrated over a segment to give an overall error score such as root mean square (RMS) error. However, critical points in the task (e.g., touchdown) can only be measured with a snapshot. The analogue of RMS error for snapshots is absolute error which is averaged over trials to obtain mean absolute error. Both RMS and mean absolute error measures, when applied to task outcome variables, are referred to as quality indicators.

The overall error measure may be further separated into <u>bias and variance</u> components. The bias component of an RMS error score gives average deviation within a segment and the variance component reflects variability about the average. The snapshot analogues are average error and variability across trials. In terms of successfully accomplishing the task goals, both bias and variable error are undesirable. However, variable error presumably reflects control difficulty which may generally be more difficult to correct than a bias tendency. In addition, an effect in which biases are equal but of opposite sign says nothing about differential performance quality, and would have to be given a low weight in assessing the impact of a factor on task performance. Thus, the variance component of an error score is generally more informative (although biases can suggest the cause of an effect). Measure categories are summarized in Table 3.

Criterion-referenced quality indicators were emphasized in the data analysis for this experiment. The preselected or <u>a priori</u> set of quality measures were: (1) RMS error of aircraft position relative to the desired flight path during the approach, hover, and descent, and (2) absolute error of the aircraft state from that desired at touchdown. In addition, inspection of data summaries suggested that measures of stick movement, aircraft activity, and bias, and variability components of task outcome would also be informative. Results of these analyses are presented only for measures showing meaningful effects that did not correlate highly with other measures. Redundancy was determined by examining correlations between measures. Generally, if a correlation between measures

was higher than 0.9, only one of them was analyzed. The measures chosen in this manner represent a posterior set since they were selected from a larger set on the basis of effect size. Although nominal significance levels used for examination of these measures were the same as those used for the *a priori* set, caution is advised in interpretation since the posterior set will have higher true alpha levels.

TABLE 3. CATEGORIZATION OF PERFORMANCE MEASURES

Location in Pilot-Aircraft Control Loop

Location	Applications	Example
Pilot Input	Control technique, workload	Collective movement
Aircraft Output	Control accuracy control technique	Aircraft roll during hover
Task Outcome	Criterion of success	Aircraft landing deviation from optimum

Measure Types

		Error Categories		
Temporal	Example	Total	Bias	Variance
Segment Summary	Lateral error during hover	RMS error	Average error	Average within-trial variability
Snapshot	Aircraft position at touchdown	Absolute error	Average error	Between-trial variability

PROCEDURES

Pilots were briefed on the use of the simulator and on the task. A document that was used in this briefing is reproduced here as Appendix A. Pilots flew 24 preliminary trials under several experimental configurations in which scene detail, field of view, and g-seat rate cuing were varied. G-seat vibration (off), collective sound (off), seastate (calm), and visual lag (117 msec), were held constant. Pilots then flew a total of 16 experimental conditions each, with four trials per condition. Pilots flew eight trials in each experimental

session, and alternated sessions with one other pilot who was going through the experimental sequence at the same time. The sessions (including preliminary trials) were more or less evenly spaced for each pilot across 2-1/2 working days. At the end of their experimental trials, pilots filled out an opinion questionnaire regarding the simulation and the experimental factors.

EXPERIMENTAL DESIGN

The experimental design was a repeated measures fractional factorial. One-half of a full 2^8 was run across the pilots such that within experimental factors, the main effects and lower-order interactions were confounded only with the high-order interactions (five-way and above) which were assumed to be negligible for the analyses. However, each pilot flew with a different set of conditions of the half fraction and this created some partial confounding of lower-order pilot-by-factor interactions with lower-order experimental factor interactions. All main effects and most two-way interactions were not confounded with any pilot-related interactions of less than the third order. The design was constructed such that main effects were generally orthogonal to any linear, quadratic, or cubic learning trends that might have been present in the data (Simon, 1977).

SECTION III

DATA ANALYSIS

The analysis of variance summaries in Tables 4 through 10 show the mean differences between levels of factors as well as η^2 (i.e., the proportion of the total variability that is associated with a specific factor effect[1]). The two-factor interactions were separated into three groups and tested omnibus fashion. For example, the seven two-factor interactions involving the pilot-experience factor were summed into the single term indicated on the tables. The total sums of squares accounted for by the sets of two-way interactions, divided by the total degrees of freedom involved, was compared to the residual mean square for tests of significance. Although these omnibus terms are shown in the tables, each estimable two-way interaction was also examined individually.

PRIMARY CONTRASTS

All effects were tested against a single residual mean square term. The residual term is composed of all sources of variance except the main effects and the two-way interactions. This includes all the three-way and higher-order interactions as well as all trial and replication effects. Thus, a large portion of the residual term estimates the within-subject, trial-to-trial variability.

Although all factors are tested against the single residual term, this is appropriate only if the pilot interaction with that factor is small. Otherwise, the main effect must be significantly larger than the interaction for the main effect to be generalizable to the population of helicopter pilots (Winer, 1971, Chapter 7). In most cases, this does not affect the interpretation of the statistical significance as presented in the tables.

TWO-WAY INTERACTIONS INVOLVING PILOT DIFFERENCES

Pilots were differentiated in the experiment on the basis of their flight experience. Thus, interactions of flight experience with other factors pose no problem for the analyses. Such an interaction requires that the factor effect in question be interpreted in relation to experience, and that

---

[1] More correctly, η_2 is the proportion of the total sums of squares of a dependent variable that is associated with group membership or designated by an independent variable.

it be generalized only to a population of pilots classified similarly in terms of experience.

Subject-by-factor interactions that are independent of experience-by-factor interactions are also possible. They would suggest the existence of a pilot classification that could impact interpretation of the relevant factor. Nevertheless, because that classification has not been identified in the experiment, it is not possible to modify interpretation of factor effects on the basis of those pilot characteristics, and the subject-by-factor interaction must be viewed as unexplained or error variance. In such a case, a factor effect must be significantly greater than the subject-by-factor interaction for it to be generalizable to the population of pilots. However, the existence of the interaction might encourage a search for the appropriate pilot classification, because its use in future experiments could extend our understanding of how other factors affect performance. This would be particularly true in an X-type interaction where the ordering of pilots in terms of performance levels is reversed between levels of a factor.

In addition, an unclassified subject-by-factor interaction, in which relationship of the relevant factor to different subjects classifications cannot be determined, will reduce the sensitivity of the statistical tests because, although larger than the residual term, it must be used as the error term in the test of significance. Only if the factor effect is statistically significant in relation to the subject-by-factor interaction can we generalize to the population of pilots from which our sample has been drawn.

In the absence of significant unclassified subject-by-factor interactions, all subject-by-factor terms may be pooled with the residual error term to strengthen the test of statistical significance by enhancing the stability of the error term (e.g., Table 8). That procedure is followed in this report wherever possible. Some cases of significant unclassified subject-by-factor interactions are presented. These are identified where they do appear, and the error term for the statistical test is adjusted accordingly by using the subject-by-factor interaction only (e.g., Table 4).

OTHER INFORMATION

There are certain computations that can be performed to obtain supplementary information with which to interpret the data. The numbers required for these analyses are available in the tables. The <u>mean performance</u> for high and low levels of any two-level factor can be obtained by taking the grand mean shown at the bottom of each column and to it add (high level) or subtract (low level) half of the mean difference for that factor. The size of the difference must be taken into consideration in the calculation since the mean of the low

condition was always subtracted from the mean of the high condition to obtain a mean difference. Thus, a negative RMS or absolute error mean difference indicates better performance with a factor's high level.

F-ratios for a particular effect can be calculated using the percentages in the table. The numerator of the ratio is the percent-variance-accounted-for by an effect divided by its degrees of freedom, and the denominator is the residual percent-variance-accounted-for divided by its degrees of freedom. Significance is indicated at .05 and .01 levels. We view .01 as the appropriate level of significance because it provides some compensation for the multiple tests and measures involved.

INTERPRETATION STANDARDS. Statistical significance, by itself, provides little information regarding the operational importance or relative size of an effect. Cohen (1977, p. 25-27) suggested, as one guideline, that effect size labels of small, moderate, and large be associated with η^2 values of 1%, 6%, and 14%, respectively. Although these values are arbitrary, we view them as reasonable guides for evaluating how much attention should be paid to a statistically significant effect.

MEASUREMENT PRIORITIZATION. For purposes of analysis, the data were categorized into several sets. First, measures from the approach, hover and descent segments, and touchdown were treated as separate groups. These three segments -- approach, hover, and descent -- are defined earlier in Table 2. Secondly, measures were grouped into *a priori* and *posterior* sets. The *a priori* set consisted of preselected quality indicators for each segment that were considered the most important in terms of describing performance outcome for a segment. These measures were generally the RMS or absolute scores for deviations from the optimum flight path although, in several cases, the variance scores were substituted when the optimum flight path or position was not well defined. This *a priori* set should be considered the more practically powerful since it involves a relatively small number of preselected scores. Detailed analyses for these scores are given in Tables 4 through 8.

All other summary measures were included in a *posterior* set. These scores may or may not involve performance quality, but can be informative and can provide secondary support for certain hypotheses. They are less powerful in an overall statistical sense because analysis involves inspection of many measures to search for effects. Tables 9 and 10 give details of analyses for certain selected posterior summary measures.

SECTION IV

RESULTS

APPROACH

The effects for line up and glideslope variability are shown in Table 4. These measures are presented rather than RMS error because the task-defined optimums did not have to be strictly adhered to. For example, the glideslope indicator guidance system which is sometimes used operationally was not functioning for this experiment. In these cases, the variability measures which indicate variability about the mean flight path within each trial are considered more appropriate measures of pilot control and quality than the RMS error scores.

The only experimental equipment effect of any consequence during the approach is that of scene detail. Average lateral variability is 11-feet greater with low scene detail. The lack of scene features (most likely the absence of a ship's wake) was apparently detrimental to pilots' ability to track smoothly toward the ship. There was also a bias effect (not shown here) of about 5 feet, which may have resulted from an attempt to line up with the side of the deck during approaches to the low-detail ship.

There were also small effects in both the lateral and vertical dimensions of the approach due to the field-of-view factor. These effects favor the wide field of view, but as they account for less than 1% of the experimental variance, they should be regarded as minor. As expected, the pilot main effects were substantial, although the pilot experience effects were inconsistent. Pilot-experience two-factor interactions were generally relatively small, but some other pilot-by-experimental factor interactions were substantial.

HOVER

The hover segment had a number of missing values, primarily because pilots often came in too low over the ramp to record time in the hover segment (see segment definitions earlier in this report). Altogether, there were no data recorded in the hover segment for 92 of 512 total trials. This created an imbalance in the experimental design and caused some nominally independent factors to be correlated. Since much of the missing data came from two pilots, induced correlations between pilot effects and experimental factors are of particular concern. This problem is compounded by the fact that pilots differed widely in average time spent in the hover segment as it was defined. Therefore, the results shown for this segment (Table 5) should be considered more suggestive than definitive, particularly with regard to small effects.

TABLE 4. ANALYSIS OF VARIANCE SUMMARIES FOR APPROACH SEGMENT VARIABILITY MEASURES

Source of Variance	LEVELS High	Low	df	Lateral (ft)	Vertical (ft)
Scene detail	high	low	1	-11.38(34.6)[a] **	0.24(0.1)
Visual lag	117 msec	217 msec	1	0.85(0.2)	-0.34(0.2)
Field of view	wide	SH-60B OFT	1	-1.58(0.6)*	-0.65(0.7)*
G-seat acceleration cuing	on	off	1	0.78(0.2)	-0.28(0.1)
G-seat vibration	on	off	1	0.73(0.2)	-0.41(0.3)
Collective sound	on	off	1	-0.24(0.0)	0.01(0.0)
Pilot experience	high	low	1	1.57(0.6)*	0.12(0.0)
Seastate/Turb.	calm	seastate	1	-1.12(0.3)	-1.62(4.2)**
Pilots			6	(9.3)**	(4.3)**
Pilot Exp. by 2-Factor Int.			7	(1.4)	(2.5)*
Other Pilot 2-Factor Int.			42	(7.9)**	(12.6)**
Other Est. 2-Factor Int.			12	(1.5)	(1.5)
Residual			434	(43.2)	(73.5)
	Grand Mean			19.41	10.02

a Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)

* $p \leq .05$

** $p \leq .01$

Table 5 gives analysis-of-variance summaries for three *a priori* hover quality measures. Lateral and longitudinal error in this segment were defined in terms of deviation of the helicopter's recovery, assisting, securing, and traversing (RAST) probe from the center of the rapid securing device (RSD) on the landing pad. The table indicates poorer performance with the low-detail scene compared to the high-detail scene, but other simulator factors do not show an effect. Note that for the pilot experience effect the more experienced pilots had less vertical variability than the less experienced pilots, but greater RMS lateral error. Thus, the pilot experience effect again shows some inconsistency.

DESCENT

Table 6 gives results for the *a priori* descent segment measures. There do not appear to be experimental effects for these outcome scores other than for pilots and turbulence. Table 7, which gives results for the lateral and longitudinal variability during descent, does show significantly increased variability under the low scene detail condition.

TOUCHDOWN

Analysis of variance summaries are given in Table 8 for touchdown quality measures. In terms of deviation from the prescribed touchdown point, touchdown was 1.21 feet closer on average under high scene detail than under low scene detail. Most of this difference was in the longitudinal dimension, although there was a significant difference in the lateral dimension also. No other simulator equipment factor affected touchdown accuracy in terms of miss distance.

Tables 9 and 10 give summaries for two other measures; vertical velocity at touchdown, and ship roll position at touchdown. Both visual lag and field of view appeared to affect vertical velocity at touchdown. For both of these factors, touchdowns were harder with the low-fidelity options. Although the field of view and visual lag effects are statistically reliable, they must be considered small since they each account for less than 1% of the variance. Furthermore, the standard deviation for vertical velocity under several conditions at sea appears to be about 1 ft/sec (Kolwey, 1977). Thus, the mean differences represent roughly one-half the standard deviation of this measure.

A visual lag effect also shows up in the measure of ship roll position at touchdown (only relevant under turbulence/seastate conditions). The average ship roll is -0.01 degrees with the shorter visual lag and 0.56 degrees with the longer visual lag. Although it is not clear why a bias effect of this nature would occur, it seems that pilots, trying to touch down with a centered ship, lag in their ability to do so in a measurable way under the longer visual lag condition.

TABLE 5. ANALYSIS OF VARIANCE SUMMARIES FOR HOVER MEASURES

Source of Variance	LEVELS High	LEVELS Low	df	RMS ERROR Longitudinal (ft)	RMS ERROR Lateral (ft)	Variability Vertical (ft)
Scene detail	high	low	1	-0.45(1.6)[a]**	-1.26(2.1)**	-0.29(1.0)**
Visual lag	117 msec	217 msec	1	-0.21(0.0)	-0.14(0.1)	-0.22(0.5)
Field of view	Wide	SH-60B OFT	1	-0.14(0.5)	-0.09(0.0)	0.00(0.3)
G-seat acceleration cuing	on	off	1	0.05(0.0)	-0.41(0.4)	0.08(0.2)
G-seat Vibration	on	off	1	-0.37(0.4)	0.16(0.3)	-0.04(0.0)
Collective sound	on	off	1	0.22(0.0)	0.31(0.4)	-0.08(0.[illegible])
Pilot Exp.	high	low	1	-0.27(0.4)	1.45(3.3)**	-0.44(2.7)**
Seastate/ turbulence	calm	seastate	1	-1.26(10.7)**	-0.03(0.1)	-0.66(9.3)**
Pilots			6	(6.8)**	(2.9)**	(22.6)**
Pilot Exp. by 2-Factor Int.			7	(0.5)	(4.6)**	(1.8)
Other Pilot 2-Factor Int.			42	(13.8)**	(9.4)	(9.8)*
Other Est. 2-Factor Int.			12	(1.6)	(4.9)*	(1.3)
Residual			344	(63.7)	(71.5)	(50.2)
	Grand Mean			2.90	9.04	1.95

a Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)

* $p \leq .05$

** $p < .01$

TABLE 6. ANALYSIS OF VARIANCE SUMMARIES FOR DESCENT SEGMENT MEASURES

Source of Variance	LEVELS High	Low	df	RMS ERROR Longitudinal (ft)	RMS ERROR Lateral (ft)	TIME Seconds
Scene detail	high	low	1	-0.10(0.1)[a]	-0.10(0.1)	-0.30(0.1)
Visual lag	117 msec	217 msec	1	-0.16(0.2)	-0.35(0.3)	-0.26(0.0)
Field of view	Wide	SH-60B OFT	1	-0.24(0.4)	0.06(0.0)	-0.79(0.3)
G-seat acceleration cuing	on	off	1	-0.18(0.2)	0.30(0.2)	-0.08(0.0)
G-seat Vibration	on	off	1	-0.20(0.3)	-0.02(0.0)	0.12(0.0)
Collective sound	on	off	1	-0.06(0.0)	0.03(0.0)	-0.85(0.4)
Pilot Exp.	high	low	1	0.00(0.0)	-0.29(0.2)	2.36(3.1)**
Seastate/ turbulence	calm	seastate	1	-0.73(3.9)**	-0.16(0.1)	-2.50(3.4)**
Pilots			6	(6.5)**	(9.9)**	(19.1)**
Pilot Exp. by 2-Factor Int.			7	(0.7)	(3.2)**	(4.4)**
Other Pilot 2-Factor Int.			42	(9.7)	(8.1)	(12.1)**
Other Est. 2-Factor Int.			12	(3.7)*	(5.5)**	(2.7)*
Residual			423	(74.2)	(72.4)	(54.4)
	Grand Mean			2.44	4.21	9.36

a Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)

* $p < .05$

** $p < .01$

TABLE 7. ANALYSIS OF VARIANCE SUMMARIES FOR SELECTED DESCENT VARIABILITY MEASURES

Source of Variance	LEVELS High	Low	df	Lateral (ft)	Longitudinal (ft)
Scene detail	high	low	1	-0.71(3.1)[a]**	-0.46(2.7)**
Visual lag	117 msec	217 msec	1	-0.04(0.0)	-0.08(0.1)
Field of view	wide	SH-60B OFT	1	0.05(0.0)	-0.11(0.2)
G-seat acceleration cuing	on	off	1	-0.06(0.0)	-0.08(0.1)
G-seat vibration	on	off	1	-0.22(0.3)	-0.07(0.1)
Collective sound	on	off	1	-0.02(0.0)	-0.10(0.2)
Pilot experience	high	low	1	0.17(0.2)	0.09(0.1)
Seastate/Turb.	calm	seastate	1	0.08(0.0)	-0.63(5.5)**
Pilots			6	(13.0)**	(3.5)**
Pilot Exp. by 2-Factor Int.			7	(1.0)	(1.4)
Other Pilot 2-Factor Int.			42	(11.1)*	(9.6)
Other Est. 2-Factor Int.			12	(3.3)	(2.4)
Residual			423	(67.8)	(73.9)
	Grand Mean			1.96	1.07

[a] Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)

* $p \leq .05$

** $p \leq .01$

TABLE 8. ANALYSIS OF VARIANCE SUMMARIES FOR TOUCHDOWN MEASURES

Source of Variance	LEVELS High	Low	df	Radial (ft)	ERROR Lateral (ft)	Longitudinal (ft)
Scene detail	high	low	1	-1.21(4.6)[a]**	-0.35(1.1)*	-1.07(3.4)**
Visual lag	117msec	217msec	1	-0.42(0.5)	-0.13(0.2)	-0.36(0.4)
Field of view	wide	SH-60B	1	-0.32(0.3)	-0.23(0.5)	-0.13(0.1)
G-seat acceleration cuing	on	off	1	0.02(0.0)	-0.21(0.4)	0.15(0.1)
G-seat vibration	on	off	1	0.11(0.0)	-0.20(0.3)	0.35(0.3)
Collective sound	on	off	1	0.03(0.0)	-0.25(0.6)	0.29(0.3)
Pilot exp.	high	low	1	-0.59(1.0)**	0.09(0.1)	-0.69(1.5)**
Seastate/ turb.	calm	seastate	1	-0.34(0.4)	-0.29(0.9)*	-0.15(0.1)
Pilots			6	(11.4)**	(6.6)**	(11.1)**
Pilot exp. by 2 factor Int.			7	(3.6)**	(1.3)	(4.0)**
Other pilot 2-factor Int.			42	(8.7)	(8.4)	(7.7)
Other est. 2 factor Int.			12	(2.4)	(3.5)	(1.8)
Residual			426	(67.0)	(76.2)	(69.3)
		Grand Mean		4.14	1.95	3.22

[a] Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)

* $p \leq .05$

** $p \leq .01$

TABLE 9. ANALYSIS OF VARIANCE SUMMARY
FOR VERTICAL VELOCITY AT TOUCHDOWN

Source of Variance	LEVELS High	Low	df	Vertical Velocity (ft/sec)
Scene detail	high	low	1	-1.67(9.8)[a] **
Visual lag	117 msec	217 msec	1	-0.44(0.7)**
Field of view	wide	SH-60B	1	-0.49(0.9)**
G-seat acceleration cuing	on	off	1	-0.01(0.)
G-seat vibration	on	off	1	-0.03(0.0)
Collective sound	on	off	1	0.05(0.0)
Pilot experience	high	low	1	-0.57(1.1)**
Seastate/turb.	calm	seastate	1	-0.71(1.8)**
Pilots			6	(28.7)**
Pilot Exp. by 2-factor Int.			7	(1.2)
Other pilot 2-factor Int.			42	(9.9)**
Other est. 2-factor Int.			12	(1.5)
Residual			426	(44.3)
		Grand Mean		5.09

[a] Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)

* $p \leq .05$

** $p \leq .01$

TABLE 10. ANALYSIS OF VARIANCE SUMMARY FOR SHIP ROLL POSITION AT TOUCHDOWN

Source of Variance	LEVELS High	Low	df	Roll (deg)
Scene detail	high	low	1	-0.11(0.0)[a]
Visual lag	117 msec	217 msec	1	-0.57(2.2)*
Field of view	wide	SH-60B OFT	1	-0.04(0.0)
G-seat acceleration cuing	on	off	1	0.00(0.0)
G-seat vibration	on	off	1	0.05(0.0)
Collective sound	on	off	1	0.13(0.1)
Pilot experience	high	low	1	0.32(0.7)
Pilots			6	(5.6)*
Pilot Exp. by 2-Factor Int.			6	(1.8)
Other Est. 2-Factor Int.			8	(1.1)
Residual			221	(88.4)
		Grand Mean		0.27

a Mean difference, i.e., mean for high level condition minus mean for low level condition. (Values of η^2 in parentheses.)
* $p \leq .05$
* $p \leq .01$

MULTIVARIATE ANALYSES

A number of multivariate analyses were made by task segment without regard to measure-type classification to assess the overall impact of the experimental effects on performance. Step-down discriminant analyses were performed for each factor on all available meaningful dependent measures from each segment. The F-to-enter criterion for a measure to be included in the discriminant function was set at 3.0 (that is, the variance accounted for by a variable above that of measures already in the equation had to represent an F of at least 3.0). At each step of the process the measure having the highest F value (at least 3.0) was added until no further measures could be added.

The strength of association between the measures and the factor can be judged in several ways. At the end of the stepping process the number of variables (measures) in the equation, the multiple correlation of the measures in the equation with the factor, and the predictability of factor group membership based on the discriminant function can be noted. BMDP program 7M (Dixon, 1981) was used to conduct the analyses and the percent correct classification values presented are jack-knifed values (Lachenbrach & Mickey, 1968). The jack-knifed procedure predicts group membership for a case using a prediction equation based on all cases except the one being predicted.

APPROACH. The discriminant results for the approach segment are given in Table 11. The bar graph shows the percent of trials correctly classified into factor group membership based on the discriminant function. The value in parentheses at the end of each bar is the number of measures that entered the equation. The next number is the multiple correlation of these measures in the equation with the factor. The candidate measures for the approach segment were: RMS, average and variable error for glideslope and line up, and stick movement scores for collective, lateral cyclic, longitudinal cyclic, and pedal.

As Table 11 indicates, the only simulator equipment factor to have any substantial multivariate effect is ship detail. Examination of the individual measures included in the function indicated that line up variability accounted for most of the variance. RMS glideslope error and line up bias (average lineup deviation within trial) also entered the discriminant function. Factor group membership (high- or low-detail conditions) was correctly predicted 79.2% of the time using this function (50% is chance with a two-group factor). The multiple correlation of the three measures with the scene-detail factor is .62 which indicates that 38% of the variance or difference in the scene-detail levels has been accounted for by the three measures.

TABLE 11. DISCRIMINANT ANALYSES SUMMARY FOR APPROACH SEGMENT

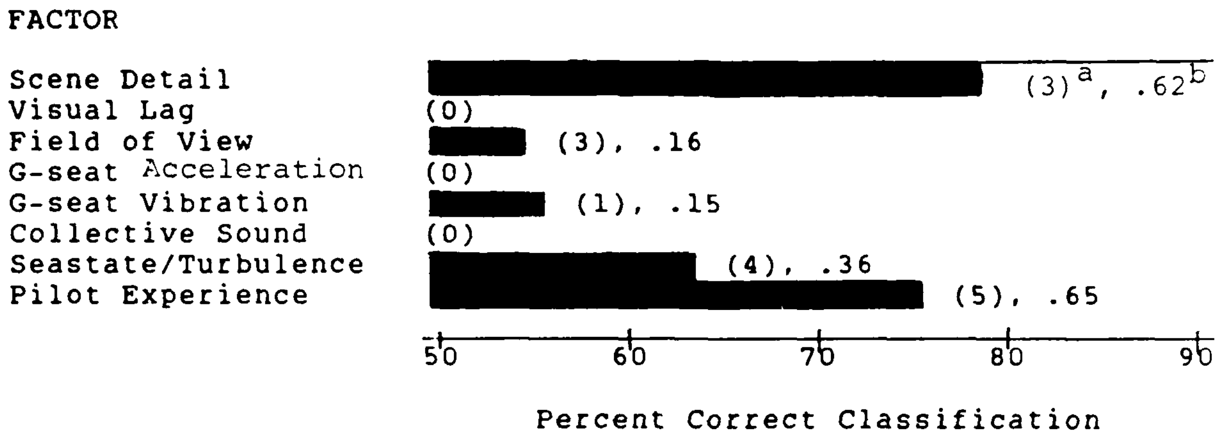
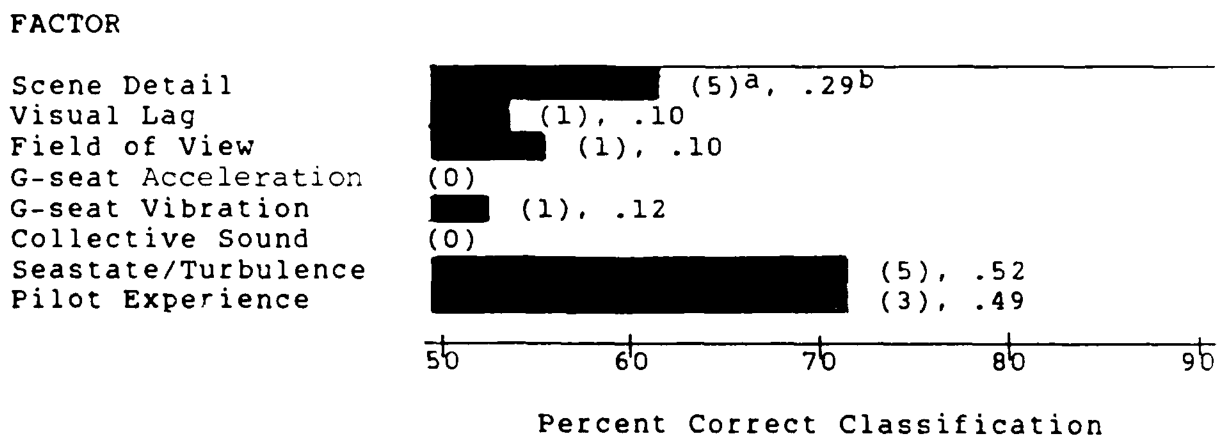


TABLE 12. DISCRIMINANT ANALYSES SUMMARY FOR HOVER SEGMENT



a Number of measures that entered the discriminant function.
b Multiple correlation of the measures that entered with the factor.

The pilot experience and seastate factors also produced substantial effects as expected. Four of the five measures in the pilot experience equation are stick movement scores. This supports previous findings (Westra et al., 1982; Westra, 1983) that pilots often differ considerably in the amount of work they do on the controls, even though they may maintain similar standards of performance.

Three measures entered the field-of-view discriminant function but they do not account for much of the experimental variance. The three measures are line up and vertical variability and RMS glideslope error. The univariate effects for line up and vertical variability are given in Table 4 and show an advantage for the wide field of view. However, the multiple correlation is .16 and group membership is correctly predicted only 54.9% of the time, barely better than chance. This discrimination is very weak.

The same holds true for the g-seat vibration factor where one measure (lateral cyclic stick movement) entered the equation. By itself, this measure is significant, showing more stick activity with the g-seat vibration cues present. No measures met the discriminant function entrance requirements for visual lag, g-seat acceleration cuing, or collective sound. Since the entrance requirements were small, it seems safe to state that these factors had little overall effect on approach performance.

HOVER. Scene detail, along with seastate and pilot experience, are the only factors with a substantial multivariate effect in the hover segment. Table 12 shows that five measures entered the discriminant function for scene detail, and that these five measures afforded group prediction success of 62.9%. Thus, the effect is not as strong as in the approach segment. (As discussed previously, there were some procedural and measurement problems which resulted in a substantial number of missing scores. It should be kept in mind that this weakens the power of analyses in this segment relative to other segments.) Measures that contributed the most in the equation were RMS lateral error, vertical bias, and lateral cyclic stick movement. In the case of vertical bias, the average hover was about 1-foot lower with the high scene detail. Lateral stick movement and RMS error were smaller with the low scene detail. Candidate measures for the hover segment, in addition to the ones mentioned, were RMS longitudinal error, average error for lateral and longitudinal performance, variable error in the vertical, lateral, and longitudinal dimensions, and the other stick movement scores.

There were no measures that met minimum requirements for the g-seat acceleration cuing factor or collective sound, while one measure each entered equations for visual lag, field of view, and g-seat vibration. The significant measure for visual lag was longitudinal variability with greater variability under

the longer lag condition. The measure entering the equation for field of view was longitudinal bias. The difference was about one-half foot with the average hover position more forward under the narrow field of view. The measure meeting entrance requirements for g-seat vibration cuing was lateral cyclic stick movement with more movement when vibration cuing was present. In a general sense, all of these effects must be regarded as very small.

DESCENT. The descent segment is the descent from hover to touchdown and a pattern of effects similar to the approach and hover segments can be seen in Table 13. The candidate measures for discriminant analyses in the descent segment were RMS, average and variable error in the lateral and longitudinal dimensions, the four stick movement measures, and time in descent. Six variables entered the discriminant function for ship detail which resulted in a function that classified 72.9% of the cases into the correct group. Measures contributing most to the function were longitudinal cyclic stick movement, collective stick movement, RMS lateral error, RMS longitudinal error, and lateral bias. RMS errors were smaller with the high scene detail while collective stick activity was also less, but longitudinal cyclic stick activity was greater.

No measures entered the discriminant functions for field of view or collective sound, while one measure each entered for visual lag, g-seat acceleration, and g-seat vibration. The measure which was significant for visual lag was lateral cyclic stick activity which was greater with the larger lag. Biases entered the equations for the g-seat factors but these were very small and afforded virtually no predictive capability for group membership (keep in mind that 50% is chance).

TOUCHDOWN. The touchdown point is not a segment, as such, but simply a capture of the touchdown point. The candidate measures for inclusion in the discriminant functions were aircraft roll and pitch, lateral, longitudinal, and velocity, and lateral and longitudinal error, all taken at the point of touchdown. Also included were the absolute value transformations of lateral error, longitudinal error, and aircraft roll, as well as the touchdown radial error from the target point. Table 14 shows a good-sized effect for ship detail along with seastate/turbulence and pilot experience but not much else. Five measures entered the equation for ship detail and the ones contributing the most were touchdown vertical velocity and touchdown radial error. Landings were harder and have greater error with low scene detail (see Table 8).

The one measure which entered the equation for visual lag was vertical velocity with harder landings taking place under the longer lag condition (see Table 9). Two variables entered the equation for field of view with aircraft pitch at touchdown contributing the most to the equation. Effects on the measures

TABLE 13. DISCRIMINANT ANALYSES
SUMMARY FOR DESCENT SEGMENT

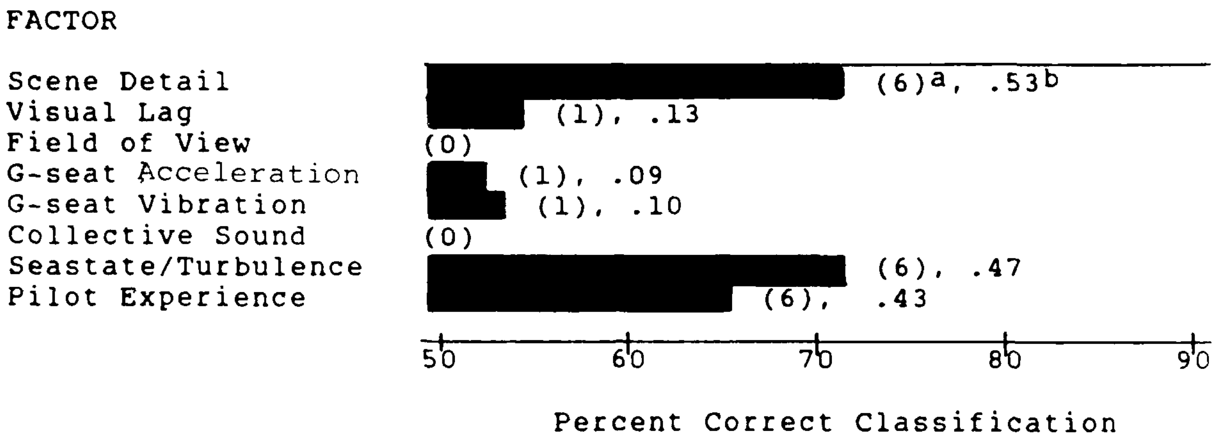
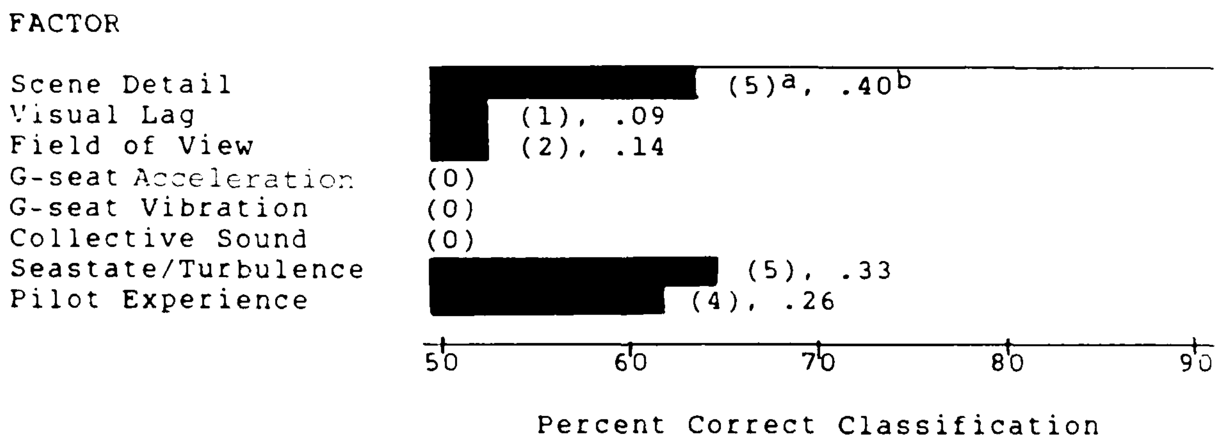


TABLE 14. DISCRIMINANT ANALYSES
SUMMARY FOR TOUCHDOWN SCORES



[a] Number of measures that entered the discriminant function.
[b] Multiple correlation of the measures that entered with the factor.

described above for both visual lag and field of view must be regarded as weak. However, analyses for ship roll and pitch position at touchdown (only relevant under turbulence/seastate conditions) did show a substantial visual lag effect for the ship roll position measure. This effect is summarized in Table 10 and discussed elsewhere. The g-seat factors and collective sound did not appear to offset touchdown in any way, not even marginally.

PILOT OPINION

Pilots filled out an opinion questionnaire following their experimental trials. A sample of the questionnaire is given in Appendix C. Pilots rated each option of the experimental factors in terms of fidelity, adequacy for training (novice pilot), and adequacy for skill retention (experienced pilots). These opinions should be considered only supplemental to the objective results as pilot opinion on some items differed considerably. Pilots were also asked to rate the perceived size of the factor effect on their performance in the simulator. As these results were not considered amenable to inferential analysis, only descriptive results are presented.

Table 15 gives the results of the pilot opinion questionnaire. In agreement with the objective results, pilots believed the g-seat factors, field of view, and collective sound had effects on performance between "none" and "small." They rated the effect of ship detail a little greater, but not large, and roughly equal to the seastate/turbulence effect. Pilots also believed the visual lag effect was between "small" and "moderate" in size, and rated it roughly equal to the ship detail and turbulence effects.

Pilots in the experiment did not like the larger lag and rated it poor in terms of fidelity and adequacy for training. They regarded it as particularly bothersome when maneuvering in close to the ship under seastate conditions. Pilots thought the g-seat vibration cuing was fairly realistic but that the g-seat acceleration cuing was distracting and unrealistic and had a negative performance effect. They preferred the simulation without g-seat acceleration cuing. They also preferred the collective sound even though they believed that it did not have a meaningful effect on performance. The SH-60B OFT field of view received fairly high ratings, although the VTRS-wide field of view was preferred.

The seastate ratings are not a matter of preference, of course, as they involve a range of real-world weather conditions rather than simulator equipment features. They are included here for comparison purposes and to indicate that pilots thought the VTRS simulation of seastate conditions was quite good.

TABLE 15. PILOT OPINION RESULTS

Factor		Fidelity	Adequacy for Training	Adequacy for Skill Retention	Estimated Effect Size
Scene detail	High	6.4[a]	6.6[a]	6.3[a]	2.3[b]
	Low	3.4	3.8	4.4	
Visual lag	117msec	6.0	6.2	6.4	2.4
	217msec	2.4	2.2	3.2	
Field of view	Wide	6.1	6.3	6.3	1.6
	SH-60B-OFT	4.2	4.2	4.8	
G-seat acceleration	On	2.3	2.3	2.9	-1.7
cuing	Off	3.3	4.1	3.9	
G-seat vibra-	On	5.1	5.2	5.0	1.7
tion	Off	2.3	3.4	3.4	
Collective	On	5.5	5.2	5.0	1.2
sound	Off	2.4	3.0	3.3	
Sea-state/	Calm	5.1	5.4	4.8	2.3
turb.	Seastate	6.1	5.8	6.4	

a Average value on a scale of 1 to 7 where "1" represents "very poor" and 7 represents "very good".

b Average value on a scale of 1 to 4 where 1 = "none", 2 = "small", 3 = "moderate", and 4 = "large".

SECTION V

DISCUSSION

The introduction to this report outlined three goals for the study, those being:

- to specify design guidelines for a simulator that could be used to teach and maintain flight skills required for SH-60B operations in proximity to a host ship.

- to screen variables for subsequent quasi-transfer and transfer experiments, and

- to develop and validate performance measures and procedures for subsequent research.

It is convenient, for the purposes of this discussion, to address those issues in reverse of the order in which they were introduced.

PERFORMANCE MEASURES AND PROCEDURES

PERFORMANCE MEASURES. In the best of circumstances, the development and validation of a performance-measurement package would result in a few measures for each of the major segments that would clearly reflect qualitative changes in pilot behavior. We could hope that these measures would be sensitive so that the interpretation of small effects would be unambiguous. In addition, we could hope for consistency, so that various manipulations that might affect a pilot's ability to perform the task would have similar effects on the performance measures or, where they did not, it would be possible to determine the reasons. While an ideal performance measurement package may not be available for any flight task, the data from this experiment suggest that the development of one for the VTOL landing task faces considerable difficulty.

Nevertheless, several specific observations can be made that will be useful for our later research with this task. Table 16 lists the measures that were influenced by either the turbulence-seastate or the pilot-experience factor. These factors can, on an *a priori* basis, be expected to impact quality of performance. Thus, the measures listed in Table 16 comprise a validated set that have been shown to reflect qualitative differences in performance.

On the positive side, Table 16 lists measures of important dimensions from all task segments. It is likely, therefore, that many qualitative changes in performance would be detected

by that set of measures. However, it was assumed that the difficulty and experience factors would have similar effects on performance, but Table 16 shows that there is little commonality in the measurement sets associated with each factor. That observation raises the possibility that some factors might affect quality of performance on task dimensions not listed in the table. Thus, the choice of those measures as our primary performance-measurement set is viewed only as a beginning, and validation for additional measures will be sought in future performance and quasi-transfer studies.

TABLE 16. PERFORMANCE MEASURES THAT SHOW VARIATION IN TASK DIFFICULTY OR PILOT EXPERIENCE

Measure	Factor	Sensitivity
Approach lateral variability*	Pilot experience	low
Approach vertical variability	Task difficulty	high
Approach control activity	Pilot experience	high
Hover lateral RMS error*	Pilot experience	high
Hover longitudinal RMS error	Task difficulty	high
Hover vertical variability	Pilot experience/ Task difficulty	low medium
Descent time*	Pilot experience/ Task difficulty	low low
Descent longitudinal RMS error	Task difficulty	medium
Descent longitudinal variability	Task difficulty	high
Touchdown radial error	Pilot experience	small
Touchdown lateral error	Task difficulty	small
Touchdown longitudinal error	Pilot experience	small
Touchdown vertical velocity	Task difficulty/ Pilot experience	small small

* Trend is the reverse of that expected under the assumption of better performance from more experienced pilots.

One other important feature of Table 16 is the rated sensitivity of the measures. That sensitivity was estimated by comparing the variance due to the difficulty or experience manipulation to that due to unidentified subject differences. Where those variations were approximately equal, the sensitivity of the measure was rated as high. If the variance due to the factor was approximately half that due to unidentified subject differences, a medium rating was assigned and a low rating was assigned where that fraction was one-quarter or smaller. Those ratings are specific to this experiment, but they should be a useful guide for similar performance experiments with similarly experienced pilots. However, the ratings may overestimate the sensitivity of the measures for a transfer or quasi-transfer experiment because we can expect inexperienced subjects who are learning a difficult task to be more variable than the moderately- and highly-experienced subjects tested in this experiment. Nevertheless, we anticipate that the sensitivity of the measures will be ranked similarly even in the face of a shift in the average rating.

One remaining concern is the reversal of performance trends with three of the measures from the direction expected on the basis of the assumed superiority of the more experienced pilots. One of those measures, time in descent, also discriminates the difficulty factor, but that trend is in the expected direction. While these anomalies are puzzling, some consideration of them may suggest useful hypotheses that could guide our future research.

For example, during discussions with pre-experimental pilots it became apparent that in the normal mode of performing the task, careful attention was paid to specific markings and features on the ship. Pilots consciously searched for visual cues that they could line up with features in the cockpit so that they could calibrate their hover and descent to touchdown. The fact that more experienced pilots spend more time in the descent segment suggests that they are more attentive and more careful in this regard.

That conclusion is supported to some extent by the observation that experienced pilots also score higher on some dimensions of control activity in different segments of the task. It is also noteworthy that this higher control activity is not generally associated with better performance on the associated measure of task outcome and on two occasions approach lateral variability/lateral cyclic activity, and hover lateral RMS error/lateral cyclic activity, is associated with poorer performance.

The observation that experienced pilots are more careful suggests an emphasis on conscious attention to control activity in this task. However, it is well known that many dimensions of control activity can be automated and are thus removed from

conscious attention (Liebowitz, Post, Brandt, and Dichgans, 1982). Control behavior under these circumstances can be more efficient, especially in high-workload situations. Schneider (1982), for example, has spoken of the need to automate behavior. He has also spoken of the difficulty that some people have of moving from a consciously attentive mode to a more automatic mode of behavior. Our data suggest that, at least on some dimensions, the greater attentiveness of the more experienced pilots resulted in poorer performance.

Thus, there is a possibility that pilots rely too heavily on inefficient conscious processes when more efficient automatic processes are potentially available. The work of Schneider (1982) would suggest that this is an inefficient means of executing a high workload task, and that pilots may be better served by reliance on automatic processes. There is some indication in our data that the more experienced pilots employ higher workload strategies which do not facilitate performance and in fact sometimes degrade it. There may be some value in future quasi-transfer experiments of evaluating instructional methods that would encourage more automatic behavior, and of comparing task proficiency of subjects who have developed automatic processes to that of subjects who rely on conscious processes to control the simulator. Further research will examine the desirability of using automatic processes where conscious processes now dominate. Techniques that might promote automatic processes, such as provision only of scene information that can be used by the ambient system, will also be examined.

PROCEDURES. In general, the procedures that were developed for the experiment appeared to have worked well. These included manipulation of variables, definition of the task, and instructions to pilots. The only obvious problem was related to the hover segment of the task. It is uncertain whether instruction to the pilots or definition of the segment caused the specific problem. It is also possible that some pilots hover for such a short time that data collection for this segment may not be practicable.

Various attempts were made to recover hover segment measures from raw data recordings, but with limited success. Attempts were also made to identify missing data with a particular experimental condition, but no strong association was found. However, almost all of the missing data involved two of the eight experimental pilots. These two pilots habitually came in low over the stern of the ship; lower than would be considered safe in normal conditions. Since the hover segment is critical from both a task and measurement point of view, it is imperative that procedures and appropriate measurement software modification be implemented for future experiments so that this problem does not occur again.

EQUIPMENT FACTORS

SCENE DETAIL. This was the most important of the equipment factors in terms of its effects on performance. Recall that the low-detail ship was depicted by solid surfaces with adjoining surfaces differing in their shades of gray. For the high-detail scene, shading demarcations between different areas and high-contrast markings were added to the deck, hangar wall, and other parts of the ship. A ship's wake and large, light blue, irregular patches on the seascape were also added. The markings enhanced the subjective impression of realism and three-dimensionality, and also provided considerably more information of the type that would enable detection of line up, position, and drift through position matching and parallax judgments. The light blue patches in the seascape appear to aid altitude perception and may have aided the ambient processes that support orientation and velocity judgments in the approach. The addition of features on the ship may have similarly aided ambient processes during close-in hover and descent to touchdown.

Many of the scene detail effects appear to result from focal and conscious processes such as position matching and judgment of parallax. Performance in terms of tracking the extended center line of the ship during the approach and of maintaining position over the desired landing point during the hover and descent to touchdown was better with the high-detail ship. As already noted, pre-experimental pilots had indicated that they looked for specific cues to maintain position and altitude in the hover, and that they lined up features on the aircraft to features on the ship. A small number of strategically placed features would provide adequate support for these types of judgments. The use of these features in this manner does not appear to offer any perceptual learning requirement so that the representation of features and markings would not necessarily have to be realistic. Judicious placement of several high-contrast patterns may suffice.

Nevertheless, it is not possible to entirely discount the role of automatic processes in perceptual judgments. Recall that scene detail had a small effect on RMS glideslope error in the approach. It is difficult to imagine how this dimension of the task could be influenced by line up or positioning of specific features. Perspective of the ship or of features on it may have affected pilots' ability to maintain the desired glideslope, or patterns in the seascape may have affected judgments of speed and altitude. Neither of these types of judgments appears to be amenable to conscious control.

In addition, the standard of performance with the low-detail ship indicates that judgments other than those of line up and feature matching that were stressed by our pre-experimental pilots can support performance of the task. While there was a noticeable loss in performance with the

low-detail ship, pilots could nevertheless execute the task and, in particular, could hover and descend to touchdown in an acceptable manner. Ambient processes that detect looming (e.g., of the hangar wall), positioning, or drift could be effective in the absence of high-contrast features on surfaces around the landing area, and may have a significant role to play in this task.

One means of distinguishing conscious and automatic processes may be available in judicious use of seastate with variations in the type of scene content that is provided. We might assume that, at least in some segments of the task, performance differences resulting from seastate would be related to perception of ship movement. Note that the presence of air turbulence could confound this interpretation so that it would be necessary to modify the procedure used in this experiment of varying air turbulence with seastate. Thus, variations in scene content could be compared with seastate active to isolate features that can influence performance. Contrasts between sets of features designed to support foveal or conscious processes versus those designed to support ambient or automatic processes should be particularly informative.

The distinction between conscious and automatic processes may have important implications for the representation of scene detail. Features designed to support conscious processes for control tasks would appear to demand less in the way of realism than features designed to support automatic processes. However, this emphasis would reverse for recognition tasks. The requirements for scene detail to support automatic processes are not well understood but realistic shape, texture, contrast, size, density, and frequency are possibilities. However, it has been established that the development of automatic processes requires extensive training and that changes in stimuli can disrupt automatic processing (Shiffrin and Schneider, 1977). Thus, there appears to be an opportunity to use simulators to teach automatic processes that may help pilots perform this task, but adequate representation of scene content is probably essential. Our future research will address that issue.

VISUAL DISPLAY LAG. This factor had some small effects on objective performance measures. This is consistent with previous work at the VTRS (Westra et al., 1981; Westra, 1982), although a considerable amount of tracking data do show that visual display lag can disrupt performance (Poulton, 1974). As noted by Poulton, subjects can apparently compensate for lags if they can anticipate the course they must follow. As a predictable course is characteristic of the flight tasks studied at the VTRS, the limited impact of visual display lag is not surprising. Nevertheless, there is an implication that pilots adjust their behavior in the presence of increased lag and this could affect behavioral strategies that would be transferred from the simulator to the aircraft.

The caution that subtle and possibly inappropriate behavioral strategies may be transferred between systems that differ in their visual display lag is not generally supported in the literature (Levine, 1953; Wightman, 1983). Nevertheless, the bulk of the data are for transfer from small to large lags which is opposite to that from a simulator with substantial lag to an aircraft. It may be important to note that Levine (1953) did include one test of transfer from large to small lags within a series of experiments in which transfer was generally in the other direction. That test was the only one to show differential transfer effects of training with different visual display lags. Thus, we recommend that visual lag be considered in future transfer experiments.

FIELD OF VIEW. As performance differences were small, our smaller field of view seems to be acceptable as a high-fidelity option. However, field of view is a major cost driver and it is possible that smaller fields of view would be useful for teaching some aspects of this task. Although our smallest field of view is probably the minimum acceptable for a full-mission simulator, more reduced fields of view may be useful for part-task trainers. Thus, smaller fields of view may be examined in future transfer work.

G-SEAT ACCELERATION. There were essentially no performance effects due to g-seat acceleration cuing. Other research has shown improved performance for a helicopter hovering task with an operational g-seat (Ricard and Parrish, 1984). The effect was not large (5% reduction in vector error) for a task which was defined as a precision hover. The task contrasts with the task in the present effort which was continuous from approach to landing with a somewhat less well defined hover taking place only while setting up for the landing. Thus, it would appear that while a g-seat can result in a small performance improvement for a precisely defined hover, the contribution to overall performance for the more operationally complete approach and landing task is negligible.

However, it must be remembered that there was a performance measurement problem with the hover segment which reduced power to detect effects in this segment. Further, the pilots who participated in the experiment felt that the g-seat was not working properly and, in fact, was somewhat disruptive in nature. Although the data do not support the opinion that the g-seat caused an actual decrement in performance, the pilots should not have had a negative reaction to it. This suggests that further human engineering development and testing should take place before any further experimental work is undertaken.

G-SEAT VIBRATION AND COLLECTIVE SOUND. These remaining factors had little or no effect on objective measures of pilot performance and must be considered as low priorities for transfer research.

DESIGN REQUIREMENTS

We recognize that a performance study of this type does not determine the most cost-effective configuration for a training simulator. Nevertheless, a system that permits a pilot to perform well should provide a training environment that is satisfactory. Until transfer data is obtained to determine the most cost-effective training system, the following design guidelines are recommended for simulators which are to be used in helicopter pilot training of the shipboard landing task.

Both objective measures and pilot opinion indicate that high-quality scene detail is important. The features in our Figure 1 high-detail scene appear to be adequate. Given a CGI system of reasonable capacity, the features for our high-detail scene could be provided at little additional cost. The particular elements considered to be important are irregularly shaped patches in the seascape, the contrast between adjoining surfaces on the ship, and the high contrast detail features on surfaces surrounding the landing area.

Pilot opinion indicated that visual lag was important, although objective performance measures showed only a slight advantage for the smaller lag. Our recommendation for a design specification is in line with pilot opinion. The smaller of our two lags (117 msec) appears to be suitable for an operational flight trainer. The smaller lag is 117 $\pm$ 17 msec and is defined by the measurement technique described by Browder (1981).

Field of view had some minor effects on objective measures and pilot opinion differential was small. As this feature is a major cost driver in visual simulation, the smaller of our two fields of view is recommended. This smaller FOV is representa tive of the FOV of the current Navy SH-60B OFT. Nevertheless, we note that it was not possible to simulate full coverage of the chin window in the VTRS. There was no indication in our data that this limitation had any adverse effect on performance, but in light of the emphasis that pilots place on the view through the chin window, it may be advisable to provide full coverage of that area until solid evidence can be accumulated to show that such coverage is unnecessary.

G-seat vibration and collective sound had no effects on objective measures, but did affect pilot preference to a small extent. If these options are relatively inexpensive they should be used because they improve face validity and pilot acceptance. In the case of the g-seat, since pilots believed there was a problem with the acceleration cuing, more research on the presentation of these cues is in order before any additional evaluation takes place.

The VTRS seastate and turbulence models were highly rated by the eight experienced participant pilots in terms of

fidelity, and the results from the performance measures indicate that this variable operated adequately as a difficulty factor in the experiment. Thus, the recommendation would be to incorporate the models in an operational flight trainer and conduct additional research using quasi-transfer experimental plans to find out how this may most advantageously be incorporated into a training system.

SECTION VI

SUMMARY

In general, results suggest that the experimental equipment feature effects can be variously categorized as large, moderate, marginal, and null. The ship detail factor had a large effect on measures taken during the approach, hover, and landing. Obviously, the lack of ship markings resulted in a more difficult simulated task as indicated by less accuracy in performance in all phases of the task. The visual system lag factor had a moderate effect on performance especially marked by aircraft roll control with turbulence present. Pilots also noticed the increased lag and believed it had a detrimental effect on their performance. Field of view had marginal effects on a few performance measures with the advantage going to a wide field of view. The two g-seat factors and collective sound had essentially no meaningful effect on performance. These results are summarized in Table 17.

TABLE 17. SUMMARY OF EFFECTS

Factor	Effect Size	Segment/Measures	Best Option*
Scene detail	Moderate/Large	All segments/most quality measures, pilot opinion	High detail
Visual lag	Small/Moderate	Hover, touchdown/ roll, pitch control, pilot opinion	117 msec
Field of view	Small	Approach, hover, touchdown/line up control, aircraft pitch	Wide FOV
G-seat vibration	Small	Approach, hover, descent/stick lateral cyclic	?
G-seat acceleration	None	-	?
Collective sound	None	-	?

* Option resulting in best simulator performance. In cases where quality measures were not affected, no determination of "best" performance was possible.

The most important effects for scene detail, visual lag, and field of view are detailed below.

- Approach (with high detail)
  - Average lateral variability smaller by 11.30 feet.

- Hover (with high scene detail)
  - Longitudinal RMS error was smaller by 0.45 feet.
  - Lateral RMS error was smaller by 1.26 feet.
  - Vertical variability was smaller by 0.29 feet.
  - Average height lower by 1 foot.
  - Lateral activity of the cyclic was higher.

- Descent (with high scene detail)
  - Lateral variability was lower by 0.71 feet.
  - Longitudinal variability (and RMS error) were lower by 0.46 feet.
  - Collective activity was less.
  - Longitudinal cyclic activity was greater.

- Touchdown (with high scene detail)
  - Radial error was smaller by 1.21 feet.
  - Lateral error was smaller by 0.35 feet.
  - Longitudinal error was smaller by 1.07 feet.
  - Vertical velocity was lower by 1.67 fps.

- Hover (with smaller lag)
  - Longitudinal variability was smaller.
- Descent (with the smaller lag)
  - Lateral activity of the cyclic was lower.
- Touchdown (with smaller lag)
  - Vertical velocity at touchdown was lower by 0.44 fps.
  - Average ship roll was smaller by 0.57° under seastate conditions.

- Approach (with the wide FOV)
  - Lateral variability was lower by 0.60 feet.
  - Vertical variability was lower by 0.65 feet.
- Hover (with the wide FOV)
  - Longitudinal error was biased backward by 0.50 feet.
- Touchdown (with the wide FOV)
  - Vertical velocity was lower by 0.49 fps.

The design options and scope of the training task studied are limited. Therefore, recommendations can only be made in light of the data presented here. A field of view similar to that currently in use with the SH-60B Operational Flight Trainer appears to be adequate. Neither field of view provided full coverage for the chin window, but we recommend that such coverage be provided. Scene detail similar to that used in our high-fidelity option should be satisfactory. Particular scene features that should be provided are the irregularly-shaped patches on the seascape, good contrast between adjoining surfaces on the ship, and high contrast features on surfaces surrounding the landing area. An average visual display lag of

117 msec is also recommended. Seastate and turbulence models should be incorporated into the system, but g-seat vibration and collective sound would be desirable only if they are found to be inexpensive options. Our g-seat acceleration cuing model appeared to be unsatisfactory, and further development is required in this area.

Transfer-of-training research is needed to support and extend these recommendations. In particular, that research could explore the distinction between scene features that support focal processes and those that support ambient processes, as well as the relative importance of conscious versus automatic processing of visual information for helicopter landings on small ships. The possibility of differential transfer effects with variations in visual system display lag, and the potential usefulness of more restricted fields of view, should also be explored. Further development of the performance measurement capability is needed for that research.

REFERENCES

Bose, E. G-seat cuing system. Akron, OH: Goodyear Aerospace Corporation, Report GER-16733, January 1980.

Browder, G. B. and Butrimas, S. V. Visual technology research simulator visual and motion cue dynamics. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN IH-326, April 1981.

Cohen, J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1977.

Dixon, W. J. (Ed.) BMDP statistical software 1981. Berkeley, CA: University of California Press, 1981.

Herndon, J. W. A Visual Technology Research Simulator for Vertical Take-off and Landing. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN IH-337, March 1982.

Kolwey, H. G. Summary of helicopter airframe testing in the shipboard environment. Presented at the 33rd Annual National Forum of the American Helicopter Society, Washington, DC: May 1977.

Lachenbach, P. and Mickey, R. M. Estimation of error rates in discriminant analysis. Technometrics, 1968, 10, 1-11.

Leibowitz, H. W., Post, R. B., Brandt, T., and Dichgans, J. Implications of recent developments in dynamic spatial orientation and visual resolution for vehicle guidance. In Wertheim, A. H., Wagenaar, W. A., and Leibowitz, H. W. (Eds.). Tutorials on Motion Perception. New York: Plenum Press, 1982.

Levine, M. Transfer of tracking performance as a function of a delay between the control and the display. Dayton, OH: Wright-Patterson AFB, Wright Air Development Center, Air Research and Development Command, USAF. WADC Technical Report 53-237, 1953.

Lintern, G. and Gopher, D. Adaptive training of perceptual motor skills: Issues, results, and future directions. International Journal of Man-Machine Studies, 1978, 10, 521-551.

Lintern, G., Wightman, D. C., and Westra, D. P. An overview of the research program at the Visual Technology Research Center. Air Force Human Resources Laboratory, Phoenix, AZ: Proceedings of the 1984 Image III Conference, 1984.

Poulton, E. C. Tracking skill of manual control. New York: Academic Press, 1974.

Ricard, G. L. and Parrish, R. V. Pilot differences and motion cuing effects on simulated helicopter hover. Human Factors, 1984, 26, 249-256.

Schneider, W. Automatic/control processing concepts and their implications for the training of skills. Champaign, IL: Psychology Department, Human Attention Research Laboratory, Report HARL-ONR-8101, April 1982.

Shiffrin, R. M. and Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 1977, 84, 127-190.

Simon, C. W. New research paradigm for applied experimental psychology: A system approach. Westlake Village, CA: Canyon Research Group, Inc., Technical Report CWS-04-77A, October 1977.

Simon, C. W. Applications of advanced experimental methodology to AWAVS training research. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 77-C-0065-1, January 1979.

Westra, D. P. Simulator design features for carrier landing. II. In-simulator transfer of training. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 81-C-0105-1, December 1982.

Westra, D. P. Simulator design features for air-to-ground bombing: I. Performance experiment I. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 81-C-0105-4, September 1983.

Westra, D. P., Simon, C. W., Collyer, S. C., and Chambers, W. S. Simulator design features for carrier landing: I. Performance experiments. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 78-C-0060-7, September 1982.

Wightman, D. C. Part-task training strategies in simulated carrier landing final approach training. Orlando, FL: Naval Training Equipment Center, Human Factors Laboratory, NAVTRAEQUIPCEN TR IH-347, 1983.

Wightman, D. C. and Lintern, G. Part-task training of tracking in manual control. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN 81-C-0105-2, September 1983.

Winer, B. J. Statistical principles in experimental design. New York, NY: McGraw-Hill, 1971.

APPENDIX A

BRIEFING. FFG7 LANDINGS WITH THE SH-60B SEA HAWK IN THE VISUAL TECHNOLOGY RESEARCH SIMULATOR

SECTION I

INTRODUCTION

Welcome to the Visual Technology Research Simulator (VTRS). This is a Naval research facility developed to study the use of simulators for teaching flight skills.

The VTRS simulates an SH-60B aircraft and consists of a full-size cockpit, a 17-foot radius spherical screen which surrounds the cockpit, and control computers which run the simulator. The cockpit controls and instruments operate just as they do in a real aircraft. A visual presentation of an FFG7 and a seascape can be projected on the screen.

Because this is a controlled experiment we will be using a special sequence and schedule. This is to assure that each person in the experiment receives the same material in exactly the same manner. Be sure to ask for clarification on any points you do not understand.

Also note that the experiment is not designed to test individuals. We want to find out how different simulator configurations affect performance. Thus, the experiment is designed to test the simulator. The report of these data will not relate performances to individual pilots. Averages and other statistical summaries will be presented but it will not be possible for readers to determine how well an individual pilot flew at any time.

SECTION II

TASK

The experimental task is to land the SH-60B Sea Hawk on an FFG7 class frigate (Figure A-1). You will be set up 1/4 mile from the ship, on a 3.5 degree glideslope with 40 knots airspeed, and in a 50 fpm descent. The FFG7 will be moving north at 10 knots.

The landing position on the deck will be demonstrated before you start the experiment. When you are shown that position, take note of the deck markings and the LSO station. This position will place your rast probe over the RSD on the deck. There are two RSDs on the deck. They are located 55 feet aft of the hangar and 5 feet left and right of the deck centerline. You should land over the RSD that is 5 feet left

of the deck centerline. The correct hover height over the landing position is about 10-14 feet above the deck. This will place your eye level at about the top of the hangar deck. The landing deck is 16 feet above sea level. If a sea state is introduced, the FFG7 will be rolling, pitching, yawing, swaying, and heaving. The landing deck will be moving accordingly.

When you are released from freeze, you are to follow a 3.5 degree glideslope (although there is no GSI available in the VTRS) to the stern of the FFG7. Smoothly reduce speed and descent rate so that you can transition to a hover above the landing area (recommended hover power is 65 to 70% torque). Descend to land on the deck when you think it is safe to do so. If a sea state has been introduced, you may have to wait for the FFG7 to reach a quiescent period. The RAST will not be used in this experiment.

Try to execute all segments of the task as smoothly as possible. Also try to maintain lineup and glideslope and to minimize yaw. Once you have been released from freeze we prefer that you complete the task without backing off to take a fresh start with an approach. We realize that, especially in the hover, it is normal practice to move away from the ship if you become uncomfortable in the hover. However, we are attempting to identify the effects of various conditions on your performance, and you may effectively disguise those effects if you fly a segment more than once in any approach.

SECTION III

INSTRUMENTS AND CONTROLS

Several of the features of the SH-60B may be new to you. Those that are important in this experiment are described below.

1. Cyclic Trim (Figure A-2). A four-way stick trim switch below the hoist switch.

2. Cyclic Trim Release Button (Figure A-2). If you depress the button, move the stick, and release the button, trim will be set for the new cyclic position.

3. Collective Trim Release (Figure A-3). Depress the button to move the collective. Release the button once the collective is in the new position. The collective is now trimmed to this new position.

4. Automatic Flight Control System. The AFCS is controlled from two panels: the AFCS control panel and the stabilator control panel. Both panels are illustrated in Figure A-4. The stabilator control panel contains all the operating controls for the stabilator. All other controls are

on the AFCS control panel, and these give the pilot the capability to engage or disengage any of the operating modes of the AFCS, the trim system, and the boost servos. The AFCS control panel also contains FAIL ADVISORY lights which should be monitored to ensure that AFCS status is correct.

The only AFCS functions used in this experiment are SAS 1, SAS 2, TRIM, BOOST, and STABILATOR.

SAS--The SAS (stability augmentation system) provides rate damping and short-term dynamic stability in pitch, roll, and yaw. Both SAS switches must be turned on for this experiment.

TRIM--The TRIM function enables the entire trim system for the cyclic, collective, and pedal control systems. The trim system maintains the controls in the desired trim position and provides the pilot with a control force gradient from the trim position. The individual trim systems can be disengaged by depressing and holding the TRIM RELEASE button on the cyclic grip (Figure A-2), the AFCS ALTITUDE RELEASE button on the collective grip (Figure A-3), and the pedal microswitches (feet on pedals). A four-way cyclic trim switch is located on the cyclic grip (Figure A-2). The TRIM function must be engaged for this experiment. (Note: There is also a friction adjustment collar on the collective lever.)

BOOST--The BOOST function activates the pilot assist servos which reduce control system forces. The BOOST system must be engaged for this experiment.

STABILATOR--The stabilator control panel engages the automatic stabilator control function. The AUTO CONTROL switch must be engaged for this experiment.

5. Altitude indicator (Figure A-5) and Mode Control Panel (Figure A-6). Depressing the HOVER switch on the mode control panel will select hover mode and activate the VHA, VDA, and Vertical Velocity Pointers. The VHA pointer indicates velocity along the heading axis and full scale is +/- 40 knots. The VDA pointer indicates side velocity and full scale is +/- 40 knots. The Vertical Velocity Pointer indicates velocity along the vertical axis and full scale is +/- 1000 fpm.

6. BDHI (Figure A-7). Depressing the TACAN switch on the mode control panel will display the tacan range (in tenths of nautical miles) in the range window and tacan bearing on the No. 2 bearing pointer and the deviation bar. Depressing the DOPPLER switch on the mode control panel will display a doppler ground speed (in tenths of knots) in the range window.

7. ROTOR RPM. 100% is recommended.

8. PARKING BRAKES. The handle is to your left when seated in the cockpit. Pull the handle up to set the parking brakes. Depress pedals all the way to release the parking brakes. It is recommended that the parking brakes be set prior to shipboard landings.

SECTION IV

EXPERIMENTAL CONDITIONS

This is a psychological experiment in which we are attempting to find out how different simulation options affect performance. The data will help us decide the areas in simulation design that should be given a high priority for development. During the experimental trials you will be flying with:

1. Field of View. Our maximum and a smaller field of view. The smaller field of view represents a three-window CRT with no chin window (similar to that found in the LAMPS MARK I simulator).

2. Visual Display. Our minimum and a longer value of delay between control inputs and response of the visual system. Such delays are inherent in visual simulators. We hope to determine the value that is small enough that it will not interfere with performance.

3. Scene Content. Our most detailed scene versus a very degraded scene. We hope to ascertain what sort of detail should be enhanced.

4. G-Seat. The g-seat will always be inflated. Sometimes it will present cues to aircraft motion and sometimes it will present vibration. At other times neither or both of these features will be presented.

5. Sound. Sometimes the sound will have cues to collective changes, and sometimes it will not.

6. Turbulence and Seastate. These will either be zero or at some level that will make your task noticeably more difficult.

Some of the conditions may seem strange to you. For example, you may be puzzled that we are interested in an impoverished visual scene when we can do so much better. However, we suspect that our best is not as good as the system should be. We are currently limited by what we can do, but can learn something about how to improve our visual scene by comparing our best level with a degraded level. In this manner we can find out what features of a scene are important and we can then work on improving our scene on those dimensions.

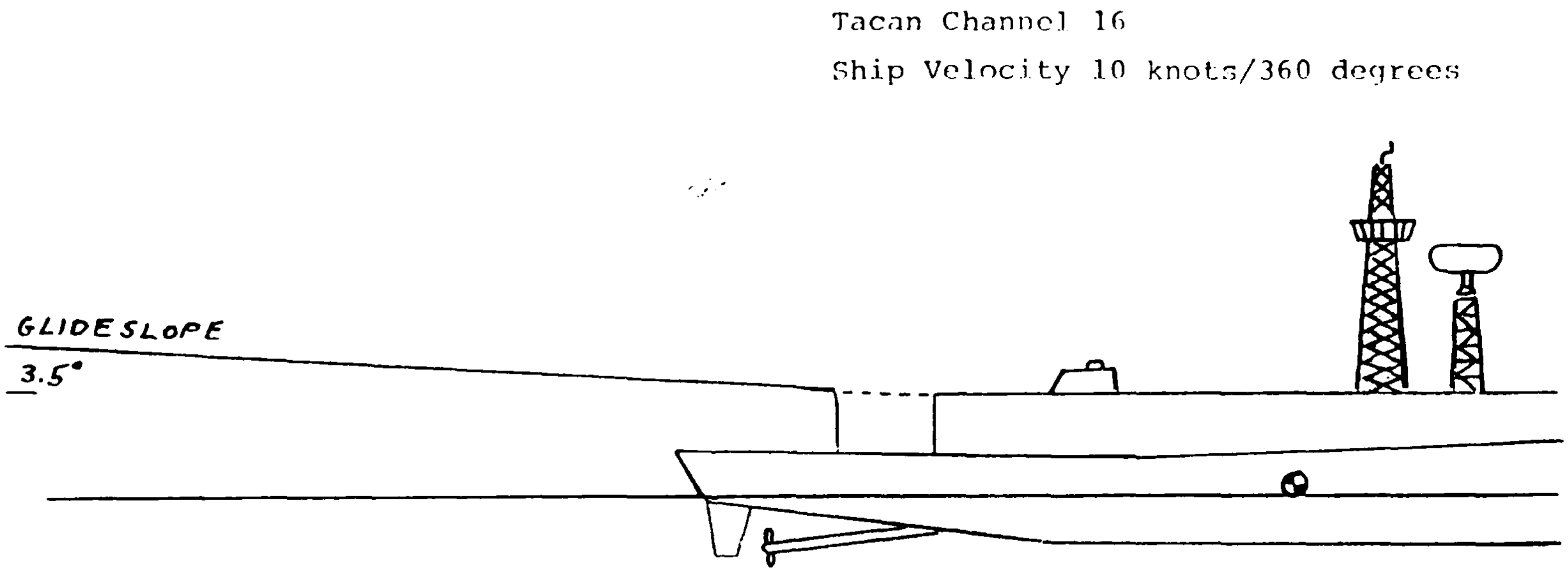


Figure A-1. Approach and Landing

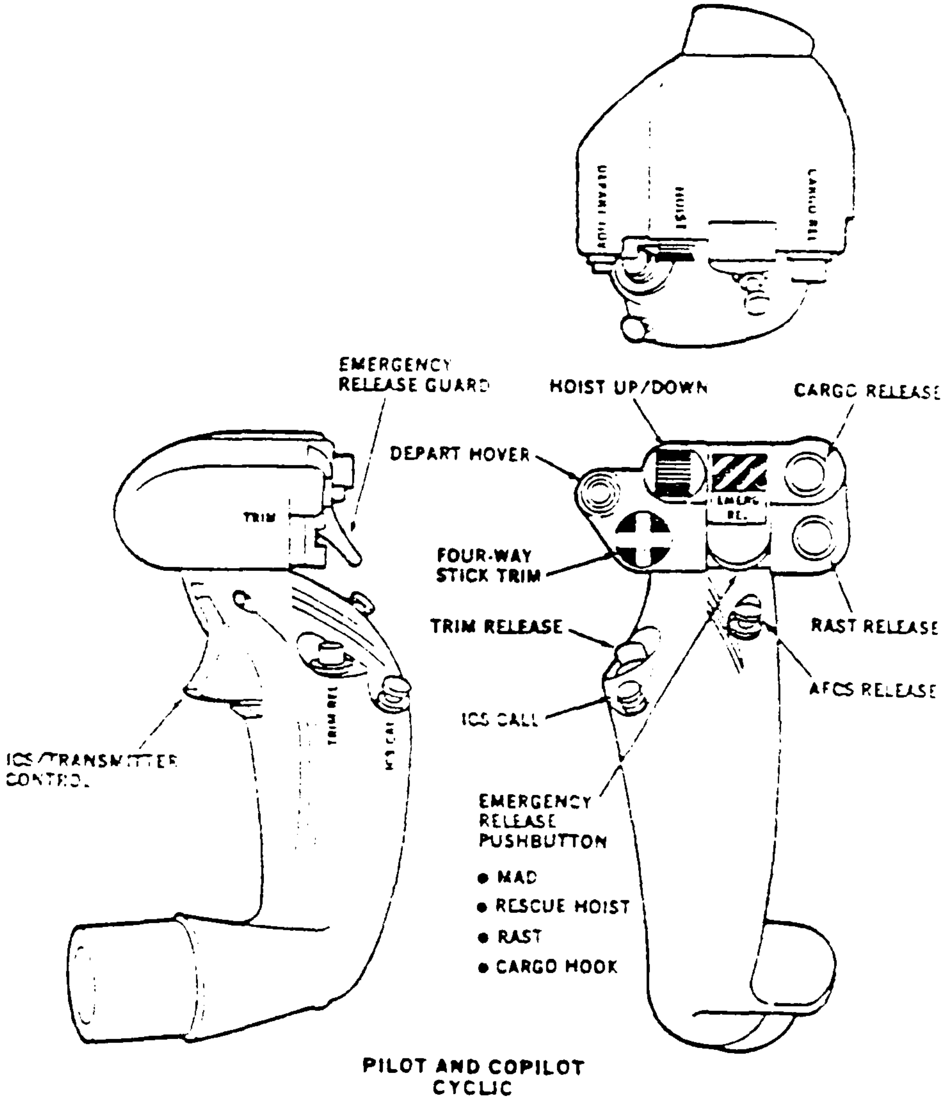


Figure A-2. Cyclic

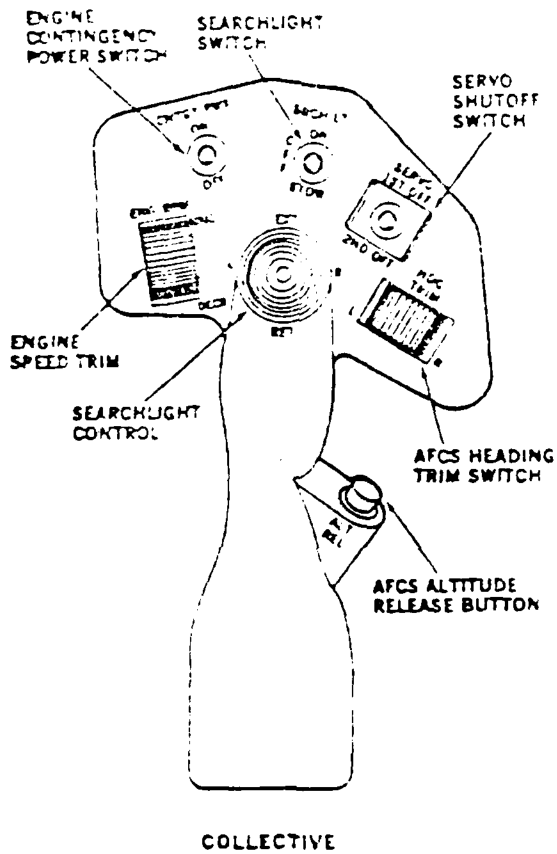
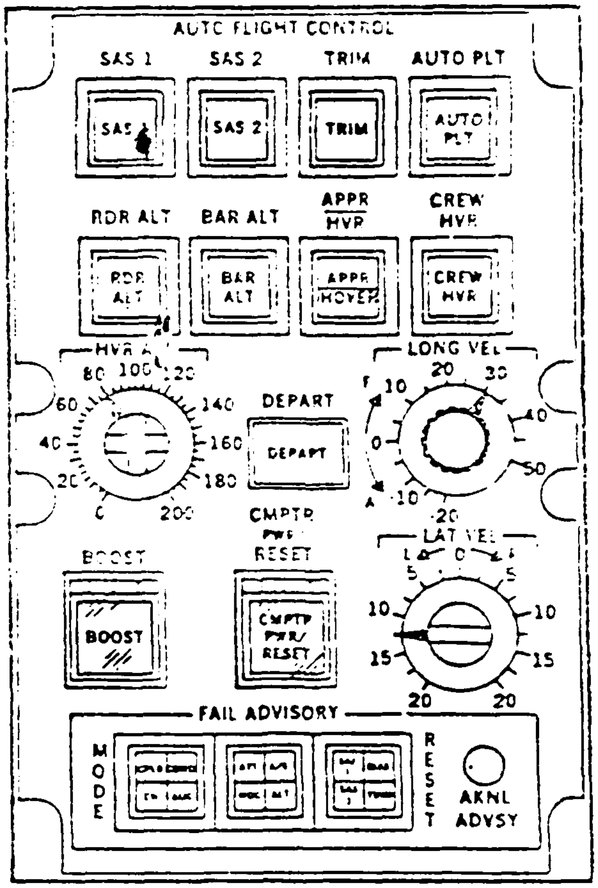
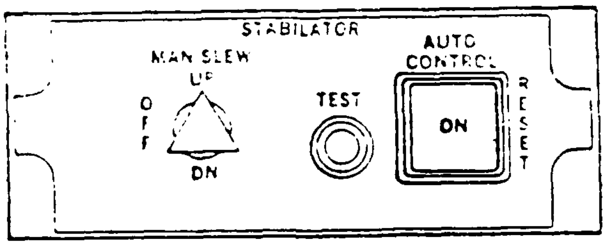


Figure A-3. Collective



AFCS Control Panel



Stabilator Control Panel

Figure A-4. AFCS and Stabilator Control Panels

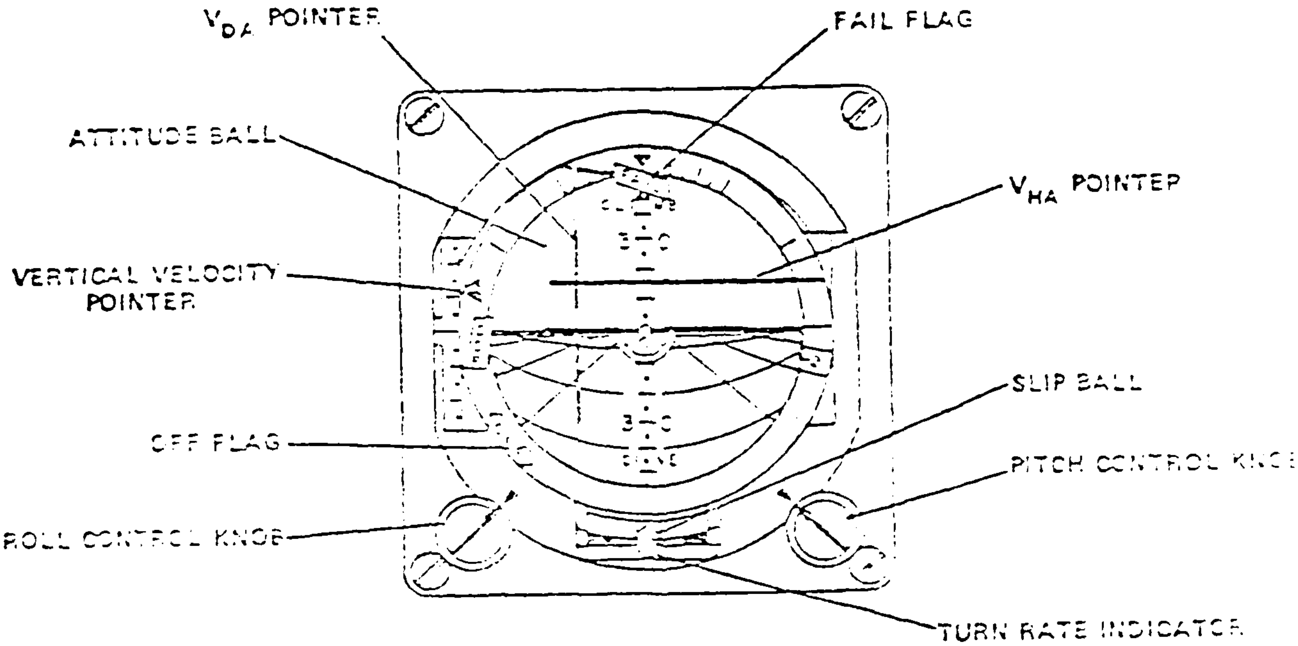


Figure A-5. Altitude Indicator

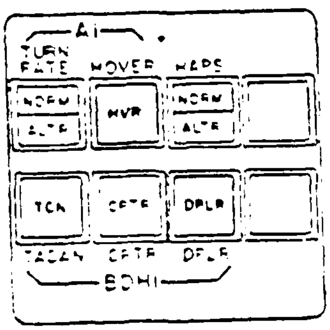


Figure A-6. Mode Control Panel

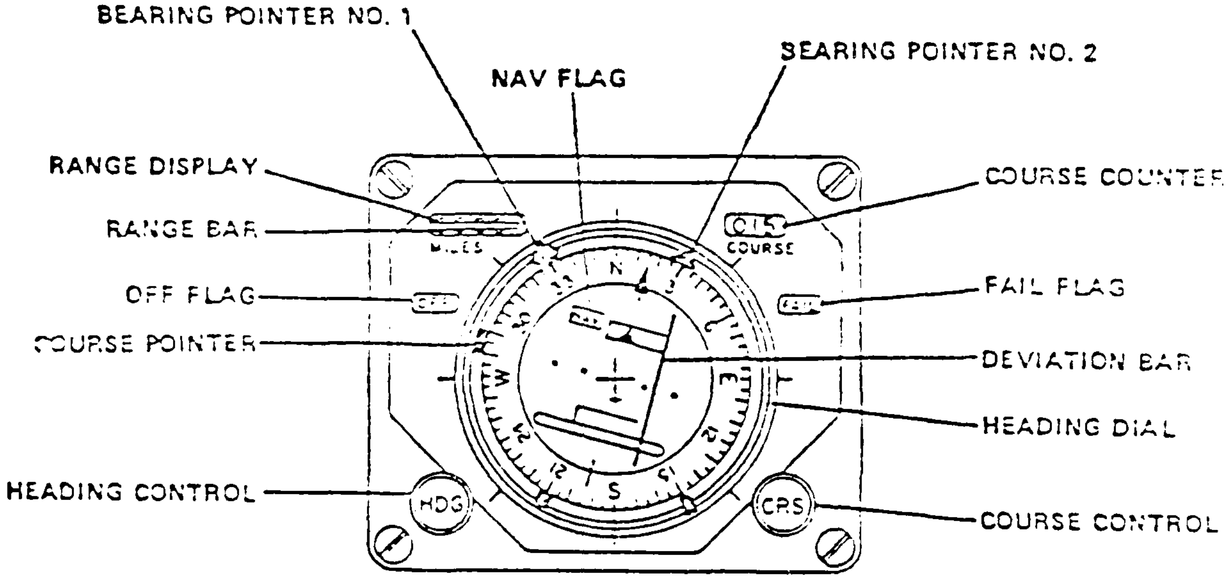


Figure A-7. Bearing Distance Heading Indicator

APPENDIX B

LIST OF PERFORMANCE MEASURES COMPUTED ON-LINE FOR EACH TRIAL

Task Segments	Variables Measured in Each Segment	Summary Measures of Each Variable
Approach	Close rate error[1] (knots)	RMS $= (\sum e_t^2/n)^{1/2}$
Hover	Lineup error[1] (knots)	Bias = e
Descent	Glideslope error (ft)	Variability $= (\sum (e-e_t)^2/n)^{1/2}$
	Glideslope error (deg)	where e_t = error at time t,
	Collective position (in)	n = number of samples
	Lateral cyclic position (in)	
	Pedal position (in)	
	Lateral error from	AST[3] $= r/n \sum(\lvert P_t - P_{t-1}\rvert)$
	touchdown point[2] (ft)	where P_t = stick position at time t,
	Longitudinal error from	r = sampling rate,
	touchdown point[2] (ft)	n = number of samples

Task Point	Variables Measured at Point
Touchdown	Aircraft roll (degrees) Aircraft pitch (degrees) Aircraft yaw (degrees) Aircraft velocity, X-axis Aircraft velocity, Y-axis Aircraft velocity, Z-axis Aircraft collective position (inches) Aircraft longitudinal error from touchdown point (feet) Aircraft lateral error from touchdown point (feet) Ship roll (degrees) Ship pitch (degrees) Ship yaw (degrees)

1 Approach segment only. Lineup error measured from landing deck centerline.
2 Hover and descent segments only. Measure is deviation of helicopter's RAST probe from the center of the RSD on the landing pad.
3 Average stick movement per second. This measure was computed only for stick variables. Other measures were not computed for stick variables.

APPENDIX C

DEBRIEF QUESTIONNAIRE

INSTRUCTIONS FOR DEBRIEF QUESTIONNAIRE

In this questionnaire, we ask you to think carefully and critically about the effect any factor investigated in the experiment may have had. For each factor, rate the options that were tested in terms of fidelity (realism), adequacy for training (novice pilot), and adequacy for skill maintenance (experienced pilot). Use a scale of 1 to 7 where "1" represents very poor and "7" represents very good.

Next, rate the size of the effect of the factor on your performance in the simulator. This should be for an "overall" performance effect. You can then further explain the nature of the effect in terms of the specific aspects of performance and the specific task segments that may have been affected. Also, try to think and comment about any combination of the factor options tested with <u>other</u> specific simulator conditions which may have particularly affected performance. For example, the g-seat vibration may have had an effect under turbulence conditions, but no effect when turbulence was not present.

The VTRS staff would like to take this opportunity to offer a special thank you for your participation in this experiment. We wish you the best in all future endeavors.

FACTOR

G-Seat Rate Cuing	Fidelity	Adequacy for Training	Adequacy for Skill Retention
ON	______	______	______
OFF	______	______	______

Difference between simulator variables on performance

EFFECT (CIRCLE ONE) Large Moderate Small None

What parameter(s) of performance?

What segments of task?

Interactions with anything else?

Other comments.

(Form for other experimental variables same as above.)

GENERAL COMMENTS

Aircraft Control Response:

Visual Scene (ship):

Visual Scene (background and seascape):

Potential of system as trainer:

Other comments:

END
DTIC
8-86